

Transactions of the ASME

The Statistical Nature of Friction	<i>E. Rabinowicz, B. G. Righmire, C. E. Tedbolm, and R. E. Williams</i>	981
The Evaluation of Corrosion Resistance for Gas-Turbine-Blade Materials	<i>W. E. Young, A. E. Hershey, and C. E. Hussey</i>	985
The Influence of Some Chemical and Physical Factors on the Formation of Deposits From Residual Fuels.	<i>P. T. Sulzer</i>	995
Compressibility Deviations for Polar Gases	<i>N. A. Hall and W. E. Ibele</i>	1003
Total Normal Emissivity Measurements on Aircraft Materials Between 100 and 800 F.	<i>N. W. Snyder, J. T. Gier, and R. V. Dunkle</i>	1011
Thermal Conductivity and Its Variability With Temperature and Pressure.	<i>L. S. Kowalczyk</i>	1021
The Ultrasonic Measurement of Hydraulic Turbine Discharge	<i>R. C. Swengel, W. B. Hess, and S. K. Waldorf</i>	1037
Recent Investigations of the Mechanics of Cavitation and Cavitation Damage	<i>Robert T. Knapp</i>	1045
On the Mechanism of Cavitation Damage	<i>M. S. Plesset and A. T. Ellis</i>	1055
Secondary Flow in Axial-Flow Turbomachinery	<i>L. H. Smith, Jr.</i>	1065
Development of a Miniature Electrohydraulic Actuator	<i>S.-Y. Lee and J. L. Shearer</i>	1077
Residual Grinding Stresses in Hardened Steel.	<i>H. R. Letner</i>	1089
The Determination of Residual Stresses in Hardened, Ground Steel	<i>L. V. Colwell, M. J. Sinnott, and J. C. Tobin</i>	1099
Temperature Distribution at the Tool-Chip Interface in Metal Cutting	<i>B. T. Chao and K. J. Trigger</i>	1107
Cutter Design and Application for Face-Milling Cast Iron and Steel	<i>O. W. Boston and W. W. Gilbert</i>	1123
Dynamics in the Inlet System of a Four-Stroke Single-Cylinder Engine	<i>C. F. Taylor, J. C. Livengood, and D. H. Tsai</i>	1133
Steam-Piping Design to Minimize Creep Concentrations	<i>Ernest L. Robinson</i>	1147
Stack Heights Required to Minimize Ground Concentrations	<i>E. W. Hewson</i>	1163

TRANSACTIONS OF THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS

VOLUME 77

OCTOBER 1955

NUMBER 7

Transactions

of The American Society of Mechanical Engineers

Published on the tenth of every month, except March, June, September, and December

OFFICERS OF THE SOCIETY:

DAVID W. R. MOROAN, *President*

JOSEPH L. KOPP, *Treasurer*

C. E. DAVIS, *Secretary*

EDGAR J. KATZ, *Asst. Treasurer*

COMMITTEE ON PUBLICATIONS:

OTTO DE LORENZII, *Chairman*

C. B. PRICE

KERR ATKINSON

W. E. REASER

JOHN DE S. COUTINHO

R. A. CORDERBERG } *Junior Advisory Members*
H. N. WEINBERG }

GEORGE A. STETSON, *Editor*

K. W. CLENDINNING, *Managing Editor*

REGIONAL ADVISORY BOARD OF THE PUBLICATIONS COMMITTEE:

RICHARD L. ANTHONY—I
JOHN DE S. COUTINHO—II
WILLIAM N. RICHARDS—III
FRANCIS C. SMITH—IV

H. M. CATHER—V
J. RUSSELL PARKER—VI
J. KENNETH SALSBURY—VII
JOHN H. KEYS—VIII

Published monthly by The American Society of Mechanical Engineers. Publication office at 20th and Northampton Streets, Easton, Pa. The editorial department is located at the headquarters of the Society, 29 West Thirty-Ninth Street, New York 18, N. Y. Cable address, "Dynamic," New York. Price \$1.50 a copy, \$12.00 annually for Transactions and the *Journal of Applied Mechanics*; to members \$1.00 a copy, \$6.00 annually. Add \$1.50 for postage to all countries outside the United States, Canada, and Pan-American Union. Changes of address must be received at Society headquarters seven weeks before they are to be effective on the mailing list. Please send old as well as new address.... By-Law: The Society shall not be responsible for statements or opinions advanced in papers or... printed in its publications (B13, Par. 4).... Entered as second-class matter March 2, 1928, at the Post Office at Easton, Pa., under the Act of August 24, 1912.... Copyrighted, 1955, by The American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the Transactions of the ASME and the author, and that date of publication be stated.

The Statistical Nature of Friction

By E. RABINOWICZ,¹ B. G. RIGHTMIRE,² C. E. TEDHOLM,³ AND R. E. WILLIAMS⁴

Sliding experiments have been carried out using copper surfaces in solid contact, and the friction traces have been analyzed statistically to study the spontaneous fluctuations in the friction force. The results suggest that the calculation of the standard deviation of the values of the instantaneous friction force can yield much information about the nature of the sliding process. High loads and smooth surface finish produce very steady traces for well-lubricated surfaces, while for unlubricated surfaces a rougher finish is required.

INTRODUCTION

IT IS well known that when sliding experiments are carried out under nonhydrodynamic conditions, using an apparatus that permits of continuous recording of the friction force, fluctuations in the instantaneous value of the friction force invariably are observed. This phenomenon has been little analyzed, it being generally assumed that the fluctuations corresponded to local differences in the surface properties of the specimens along the sliding track. When this view is examined critically, however, it is not at all obvious that surfaces cannot readily be prepared whose properties are very nearly uniform over rather large regions (say, a few square millimeters). Following a suggestion of Nagasu (1),⁵ we may more plausibly assume that the fluctuations are an essential part of the friction process. This hypothesis is explicable in the light of experiments suggesting that, during sliding, bridges are formed between the sliding metals, relatively small in number and large in size. During the sliding process these junctions have to be sheared and the force to shear them constitutes the friction force. As junctions continue to be made and broken, fluctuations in the friction force might be expected.

A series of sliding experiments have been carried out, and the friction results have been analyzed statistically in an attempt to correlate the fluctuations in the friction force with the known variables of the sliding process.

EXPERIMENTAL

Preliminary experiments showed that the fluctuations were quite small for carefully prepared surfaces, a few per cent or so, and hence the friction coefficient had to be measured with considerable precision. The apparatus used is shown in Fig. 1 and employs a $\frac{1}{8}$ -in.-diam rider having a hemispherically shaped end which contacts a flat specimen mounted on a steel turntable. The rider, which is loaded by a dead weight, is held by a support-

ing arm that is free to deflect with the frictional forces acting on the rider and, in turn, these frictional forces act on a strain gage. Both the change in resistance of the strain gage, due to the friction force, and the change in resistance of an interrupter mechanism on the rotating disk are displayed on a Sanborn two-channel recorder. The slider was cleaned on a fine emery paper while it was

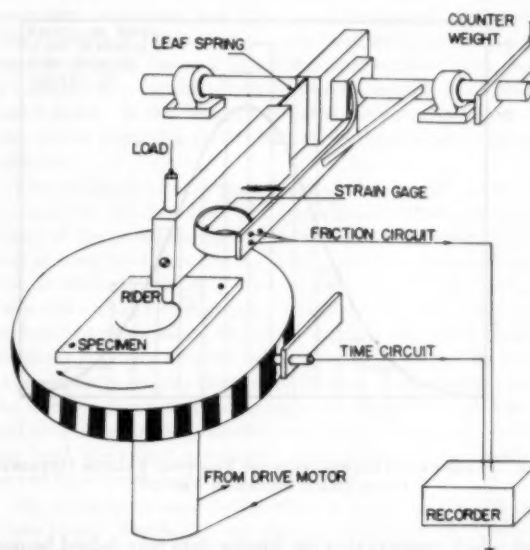


FIG. 1 FRICTION APPARATUS USED IN INVESTIGATION

rotated in the chuck of a drill press. The flat was cleaned by lapping on emery paper of the required mesh size, and tapped lightly to remove any loose emery grains from the surface. A sliding speed of about 10^{-2} cm/sec was used so as to avoid the complicating effects of frictional heating.

Readings of the friction force could be made with a probable error of about $\frac{1}{10}$ of one division, approximately $\frac{1}{2}$ per cent or 0.3 g, whichever was the higher. This sensitivity was satisfactory in almost all cases.

The statistical analysis was performed on 0.5 cm of sliding track. A section of the friction trace, of the calculated requisite length, and as free as possible from gross fluctuations corresponding to surface abnormalities, was selected and measurement of the friction force was made at the point where the friction trace crossed each successive vertical line on the recording paper. About 150 to 200 readings were taken for each trace and an average and standard deviation calculated.

RESULTS

Fig. 2 shows a typical friction trace obtained in the experiments, and Fig. 3 shows a corresponding frequency distribution of the actual friction values, suitably grouped. The continuous curve in Fig. 3 represents the corresponding normal distribution curve as calculated. Applying a χ^2 -test, in comparison of the distribution, actual and calculated, of Fig. 3, a value of P equal to 0.14 is ob-

¹ Staff Member, Division of Industrial Cooperation, Massachusetts Institute of Technology, Cambridge, Mass.

² Associate Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

³ U. S. Naval Shipyard, San Francisco, Calif.

⁴ U. S. Coast Guard Shipyard, Curtis Bay, Md.

⁵ Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Lubrication under the auspices of the Lubrication Activity of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the First Annual ASME-ASLE Lubrication Conference, Baltimore, Md., October 18-20, 1954.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 28, 1954. Paper No. 54-Lub-2.

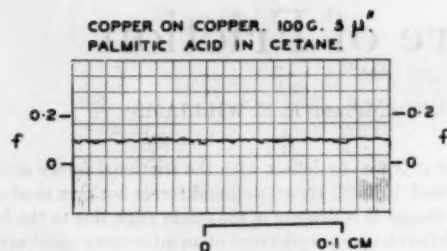


FIG. 2 TYPICAL FRICTION TRACE SHOWING FLUCTUATIONS IN THE FRICTION FORCE

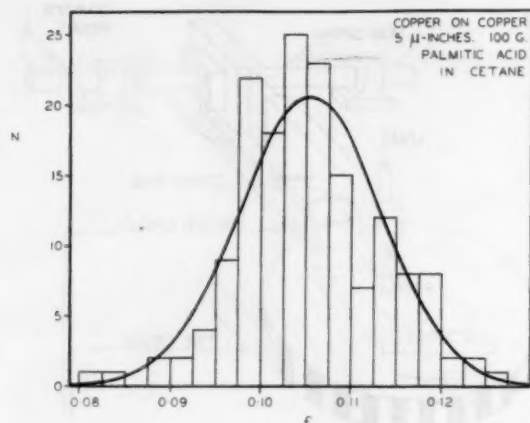


FIG. 3 FREQUENCY DISTRIBUTION OF FRICTION VALUES OBTAINED FROM TRACE SHOWN IN FIG. 2

tained, which suggests that the friction data may indeed be normally distributed about an average value.*

Often the measured readings show a more pronounced skewness, and P -values as low as 0.07 and 0.05 are obtained. However, it was felt that this did not too greatly affect the validity of our analysis.

For comparative purposes not σ the standard deviation itself, but the quantity σ divided by \bar{f} , the average friction coefficient, was used.

EFFECT OF ROUGHNESS

Fig. 4 shows results for a series of runs of copper on copper, load 500 g, lubricated with solid graphite. It will be seen that σ/f increased with increasing roughness, suggesting that with rough surfaces the rider goes upward and downward following the asperities and depressions of the surface. However, the results suggest that even with the smoothest surfaces there is a finite σ/f component.

We may calculate the importance of the roughness term from Fig. 4. If we take a surface with asperities of angle θ , then the friction coefficient will be $\bar{f} + \tan \theta$ going up the asperities and $\bar{f} - \tan \theta$ going down. Hence we obtain a σ/f component equal to $\tan \theta/f$. From Fig. 4 we see that at a roughness of 65 microinches a value of 9 per cent is obtained for σ/f , 2 per cent

* P in this calculation represents the probability that a random experiment would produce poorer agreement than that observed between the actual and normally distributed values of N as a function of f which are shown in Fig. 3. According to Fisher (2), "If P is between 0.1 and 0.9, there is certainly no reason to suspect the hypothesis tested."

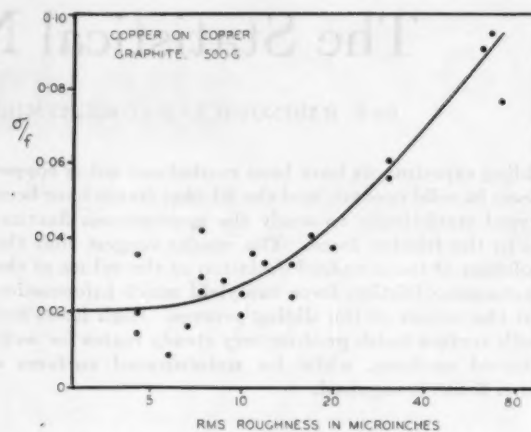


FIG. 4 PLOT OF RELATIVE STANDARD DEVIATION AS A FUNCTION OF ROUGHNESS FOR COPPER SURFACES, LOAD 500 g, LUBRICATED BY GRAPHITE

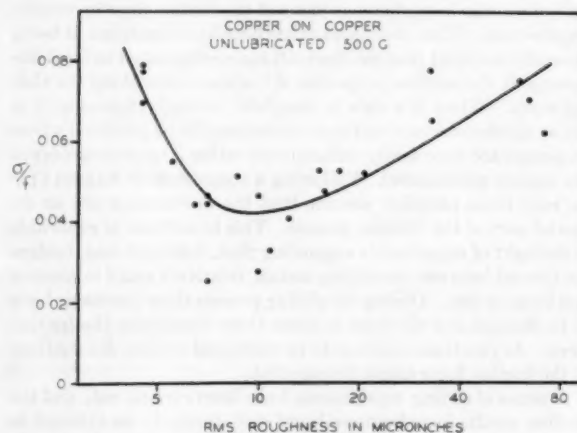


FIG. 5 PLOT OF RELATIVE STANDARD DEVIATION AS A FUNCTION OF ROUGHNESS FOR UNLUBRICATED COPPER SURFACES, LOAD 500 g

of which appears not to be due to roughness and hence must be subtracted. The remaining 7 per cent represents the roughness effect. This result is nearly the same as that of Strang and Lewis (3), who used a widely different experimental arrangement and measured the rise and fall of a rider during sliding. In any case we may conclude that only a few per cent of the total frictional force arises from a roughness effect.

Fig. 5 shows a similar experiment carried out with unlubricated copper. For very rough surfaces the results are similar to those obtained with lubricated surfaces, but for very smooth surfaces high values of σ/f are obtained. These results may be explained on the assumption that in the case of clean copper on copper, as sliding proceeds, the junctions tend to increase by a "snowballing" mechanism (4), and when one of these large junctions finally breaks off, a large drop in friction takes place. As the sliding surfaces are made rougher, the junctions are forced to break off before reaching very large dimensions, and thus smoother friction traces are obtained. In confirmation of this assumption it may be noted that for this series of experiments the coefficient of friction was not independent of roughness, but varied from 0.6 for the roughest surfaces to 1.2 for the smoothest (5), suggesting that welding was more severe in the latter case.

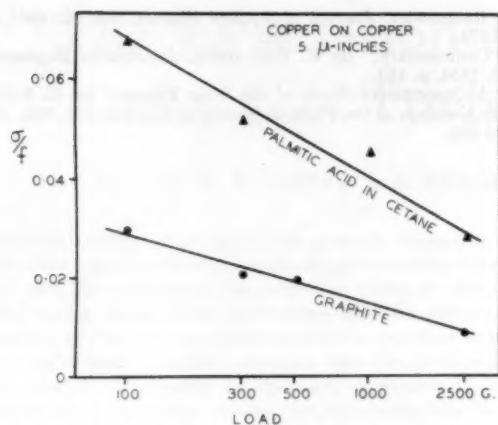


FIG. 6 PLOT OF RELATIVE STANDARD DEVIATION AS A FUNCTION OF LOAD FOR COPPER SURFACES LUBRICATED BY GRAPHITE AND BY A SOLUTION OF PALMITIC ACID IN CETANE (RMS roughness 5 microinches.)

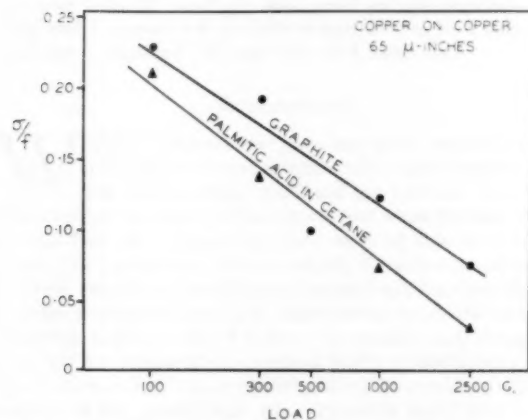


FIG. 7 PLOT OF RELATIVE STANDARD DEVIATION AS A FUNCTION OF LOAD FOR COPPER SURFACES LUBRICATED BY GRAPHITE AND BY A SOLUTION OF PALMITIC ACID IN CETANE (RMS roughness 65 microinches.)

EFFECT OF LOAD

Figs. 6 and 7 show results of experiments at varying load of copper on copper, lubricated with graphite and a solution of palmitic acid in cetane. Both for smooth and rough surfaces σ/f diminishes as the load is increased. Two independent phenomena seem at work. For very rough surfaces there is a smoothing-out effect, and the rider can follow the surface irregularities less closely at high loads; thus the roughness effect becomes less important at high loads. With smooth surfaces the large area of contact at the higher loads leads to a greater number of junctions, and this, as will be shown later, reduces the percentage deviation of the value of the friction coefficient.

DISCUSSION

The results obtained in the experiments just described show that measurement of the σ/f factor gives valuable information about the sliding conditions not provided by the coefficient of friction itself. Hence additional information may be obtained on the mechanism of friction.

When the roughness of the surfaces is increased, σ/f increases

also, and the extent of its increase shows directly the importance of the roughness component in the coefficient of friction. In the extreme case investigated of a well-lubricated surface, a roughness of 65 microinches and a load of 100 g, some 22 per cent of the sliding force was needed to overcome the roughness along the sliding track. At loads of 2500 g and with surfaces of roughness 5 microinches, less than 1 per cent of the friction force is expended in this way, and the results may be considered confirmation that the so-called roughness mechanism cannot explain more than a small percentage of observed friction values.

The fact that even the smoothest surfaces show some σ/f may be explained by the welding hypothesis. The metal surfaces in contact are assumed to form small bridges at the tips of their prominent asperities, and these must be sheared during sliding. The strength of the bridges is not a constant quantity but varies considerably, in the case of unlubricated surfaces in air, because a junction may consist of oxide-to-oxide contacts or of metal-to-metal welds; in the case of lubricated surfaces, because in addition to the foregoing a thin film of lubricant may separate the surfaces.

The condition of each junction may be assumed not to be influenced by the condition of its neighbors; thus at one instant many of the junctions may have metal-to-metal contact (f high) and at some later time the majority may have a separating oxide film or lubricating layer (f low). As sliding proceeds, the friction force will continue to fluctuate. In accordance with a well-known probability relationship, the σ/f to be expected under these conditions is inversely proportional to the square root of the number of junctions (assuming them equal in size). Now it is known that the number of junctions is increased as the load is increased (6), and thus an accompanying decrease in σ/f is to be expected. As has been mentioned previously, such a decrease is actually observed for smooth surfaces.

For rough surfaces a similar effect is noted, but in this case it seems largely due to the smoothing-out effect at the higher load produced by the larger area of contact, which does not allow the rider to follow too closely each depression and elevation on the flat surface.

The results with clean copper surfaces support the snowballing mechanism. As sliding continues, a few of the metallic junctions grow to a very large size, and as each of them is finally broken, a large decrease in f takes place, and hence large values of σ/f are obtained. As the surfaces are made rougher, this growth of the junctions is made impossible and hence smaller values of σ/f obtained.

CONCLUSION

The experimental analysis described in the foregoing has been concerned with an analysis of friction values considered as independent observations of essentially the same quantity. A different approach is to take the successive values as constituting a time series, and statistical analysis along these lines is in progress.

The engineering significance of the investigation is that it discloses those variables in the sliding process that are conducive to steady, nonfluctuating friction coefficients. In particular, the results suggest that at very light loads large fluctuations are to be expected; furthermore, while for well-lubricated surfaces the fluctuations decrease as the surface roughness is decreased, for poorly lubricated or unlubricated surfaces there is a critical range of surface roughness, with smoother and rougher surfaces alike producing much larger fluctuations in the friction force.

BIBLIOGRAPHY

- 1 "Statistical Features in Static Friction," by H. Nagasu, *Journal of the Physical Society of Japan*, vol. 6, 1951, pp. 123-124.

2 "Statistical Methods for Research Workers," by R. A. Fisher, Hafner Publishing Company, New York, N. Y., eleventh edition, 1950, p. 80.

3 "On the Magnitude of the Mechanical Component of Solid Friction," by C. D. Strang and C. R. Lewis, *Journal of Applied Physics*, vol. 20, 1949, pp. 1164-1167.

4 "The Nature of the Static and Kinetic Coefficients of Friction,"

by E. Rabinowicz, *Journal of Applied Physics*, vol. 22, 1951, pp. 1373-1379.

5 "Commentary," by E. Rabinowicz, *Lubrication Engineering*, vol. 10, 1954, p. 189.

6 "A Quantitative Study of the Wear Process," by E. Rabinowicz, *Proceedings of the Physical Society of London*, vol. 66B, 1953, pp. 929-936.



Fig. 1. Log-log plot of wear rate and friction coefficient vs. velocity. The upper line is the wear rate and the lower line is the friction coefficient. Both lines show a positive linear relationship.



Fig. 2. Log-log plot of wear rate and friction coefficient vs. velocity. The upper line is the wear rate and the lower line is the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The wear rate and friction coefficient are plotted against velocity on a log-log scale. The upper line represents the wear rate and the lower line represents the friction coefficient. Both lines show a positive linear relationship.

The Evaluation of Corrosion Resistance for Gas-Turbine-Blade Materials

By W. E. YOUNG,¹ A. E. HERSHEY,² AND C. E. HUSSEY,³ E. PITTSBURGH, PA.

Surface analysis of gas-turbine-blade material before and after exposure to the combustion products of residual fuel oils, demonstrates the corrosive effect of these products, and a measurable indication may be obtained in a fraction of the testing time required to produce appreciable weight loss. Surface analysis has the further advantage that it is possible to measure corrosion in specific regions on a specimen where the specimen has been exposed to a gas-flow pattern with a well-defined temperature profile. Weight-loss measurements following long-time tests have shown good agreement with this method of analysis. The fusion temperature of the fuel ash appears to be a reliable criterion for judging the potential corrosivity of residual fuel oil, and good correlation has been obtained between ash-fusion temperatures and corrosion both for untreated oils and oils with additives.

INTRODUCTION

WITHIN the past few years operating experience with steam generators at progressively higher temperatures, and more recently with land gas turbines, has shown that a serious corrosion problem may exist when burning certain residual fuel oils. Apparently the sources of this corrosion are certain ash constituents in the products of combustion of the oil, of which vanadium pentoxide and sodium sulphate are thought to be the worst offenders. The exact mechanism of this corrosive attack has been the subject of much speculation and, though the process is not completely understood, it seems reasonably certain that in its so-called "catastrophic" phase it is associated with the presence of the pentoxide or the sulphate in liquid form at the corroded surface.

Following is a report on a new method for obtaining quantitative measurements of corrosion after relatively short exposures (2 hr) to combustion products of residual fuel oil. Since the fusion of the fuel ash appears to be an important factor in the corrosion process, a microfusion apparatus is described and results of such fusion measurements are correlated with corrosion measurements from short-time tests and with weight-loss measurements from longer corrosion tests.

The corrosive attack associated with fuel oil having high concentrations of vanadium has been discussed frequently during the past few years (1 to 8).⁴ Hence this aspect of the problem

will be considered only briefly in this paper while major emphasis will be placed on the results obtained with oils having high sodium concentrations. The corrosive attacks from these two sources have much in common, but they also have characteristic differences such as the temperature range in which they become excessive and the relative severity of the attack on some of the well-known refractory alloys.

TEST FACILITIES AND PROCEDURE

Various test methods have been employed in attacking the problem of residual fuel-oil corrosion but, in general, they may be classified into the following types:

- 1 Crucible tests with metal samples heated in direct contact with fuel ash or ash components (3, 4).
- 2 Small furnace combustion tests with samples placed in the furnace exhaust-gas stream (4).
- 3 Gas-turbine combustion-rig tests operating at atmospheric pressure and simulated turbine operating conditions (8).
- 4 Gas-turbine combustion-rig tests operating at actual turbine pressures and gas flows (8, 9).

All of these approach the ultimate of actual use in a turbine to varying degrees at the expense of increasing complexity of the test procedure.

The corrosion results being presented here were all obtained with the equipment shown schematically in Fig. 1, operating

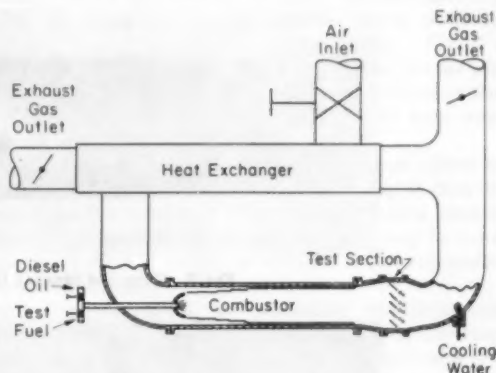


FIG. 1 SCHEMATIC PASSAGE LAYOUT

either as type 3 or 4 test rig. The combustion section is similar to that used in the 1800-hp gas turbine with a corrosion test section and water-cooling section located at the combustor outlet. The test section was designed to hold five strip specimens $2\frac{1}{4}$ in. \times $1\frac{1}{4}$ in. \times $\frac{1}{8}$ in. mounted at a 45-deg angle to the direction of gas flow. They were placed on $\frac{3}{4}$ -in. centers and, except for their uniform thickness, corresponded closely to a first-row stationary cascade. Fig. 2 is a photograph of the test section showing 16 thermocouples located in the wakes of the specimens. Fig. 3 is a photograph of the combustor tube after some 12 hr of operation at a nominal outlet temperature of 1600 F. At this

¹ Thermodynamics Section, Research Laboratories, Westinghouse Electric Corporation.

² Advisory Engineer, Research Laboratories, Westinghouse Electric Corporation. Mem. ASME.

³ Design Engineer, Steam Division, Westinghouse Electric Corporation. Assoc. Mem. ASME.

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Gas Turbine Power Division and presented at a joint session of the Gas Turbine Power and Fuels Divisions at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 15, 1954. Paper No. 54-A-215.

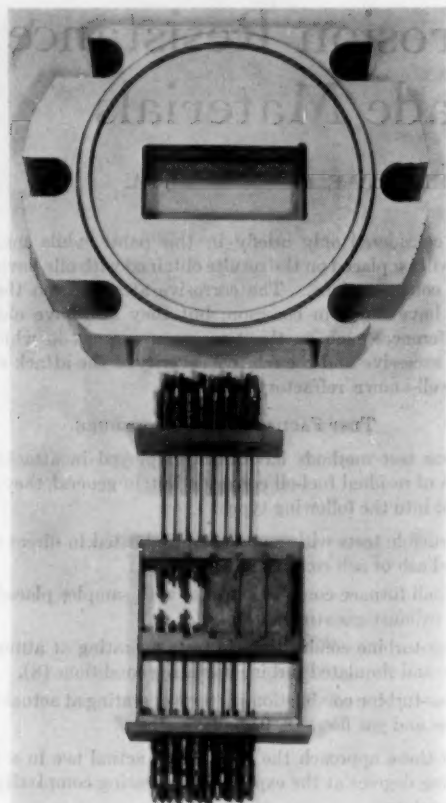


Fig. 2 SECTION WITH SPECIMEN HOLDER REMOVED

time the tube walls in the vicinity of the primary combustion zone showed only slight oxide coloring and no deformation.

This combustion equipment has been used satisfactorily in conjunction with a corrosion test program which has been in progress for the past two years. As previously stated, it has been operated at two different test conditions as indicated in Table 1. Tests at atmospheric pressure were from 40 to 100 hr in

TABLE 1 TEST CONDITIONS

	Condition 1	Condition 2
Combustor pressure, atm.	1	4
Fuel rate, lb/hr.	18	100
Air rate, lb/sec.	0.30	1.75
Velocity over specimen, fpe.	450	625

length and generally were used to verify results previously obtained with high-pressure tests of 2 hr duration. The extent of the corrosive attack during a long-time test was determined by the loss of weight of the specimen. However, during the short-time tests, the weight loss of most refractory metal specimens was too small to give significant results so that a new technique was developed, which depended on measuring the change in roughness of the surface of the specimen.

The surface-roughness measurements were made with a Brush surface analyzer with certain additional equipment as shown in Fig. 4. Since the conventional drive merely oscillates over a total length of $1/8$ in., an auxiliary drive was added making it possible to traverse the full length of the specimen. This auxiliary drive was arranged to run in a direction normal to the standard drive so that the actual path traversed was a saw-tooth wave of approximately $1/16$ -in. amplitude and $1/8$ -in. wave length. In addition to the usual amplified record of the stylus movement, an RMS meter is also in the circuit to give a visual indication of the maximum and minimum signals. However, in the present case, a continuous record of RMS readings was considered necessary since a curve of this sort is much easier to analyze than the very irregular direct surface record. Accord-

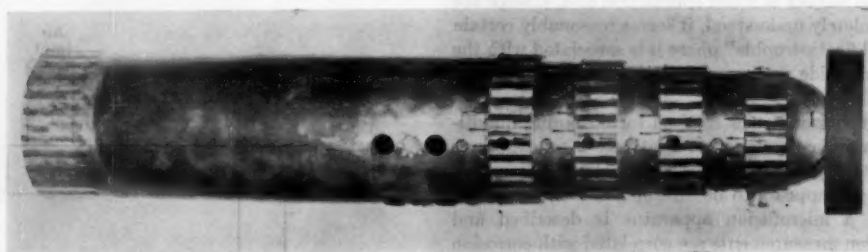


Fig. 3 TYPE B-4 SPECIAL HIGH-TEMPERATURE COMBUSTION BASKET



Fig. 4 SURFACE-ANALYZING EQUIPMENT

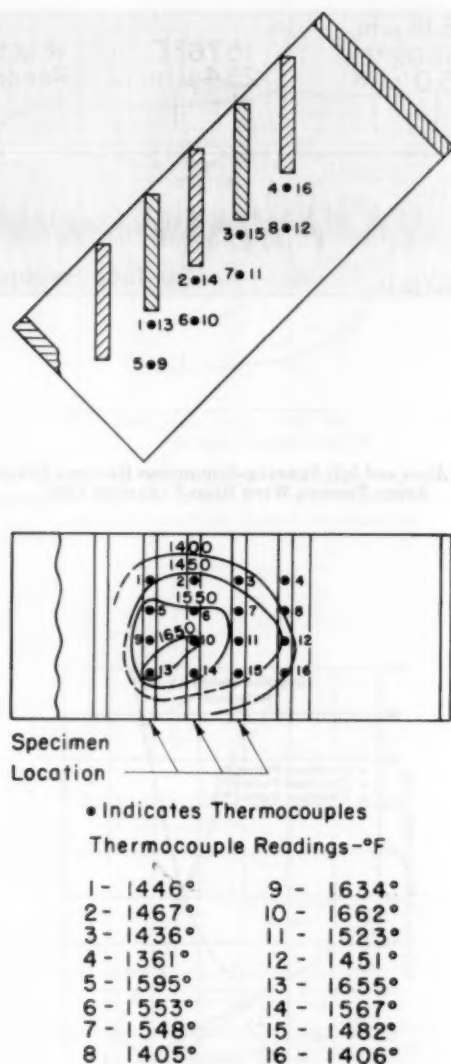


FIG. 5 TYPICAL TEMPERATURE PROFILE AT TEST SECTION

ingly, an amplified signal from the RMS meter was recorded side by side with the actual surface record by means of a double-channel oscillograph. A typical record is shown in Fig. 6.

This method has the advantage that measurable indications of corrosion can be obtained in a fraction of the testing time required to produce appreciable weight loss. And further, it is possible to measure corrosion in specific regions on a specimen which has been exposed to a gas-flow pattern with a fixed and well-defined temperature profile. It also should be possible to examine complicated sections and parts of an actual machine where a weight-loss measurement would be obscured by the large mass of the section. The limitation of the method lies in the fact that some of the original surface must be retained as a reference after exposure. If all of the original surface has been removed, it is conceivable that corrosion could proceed at an essentially constant roughness until the entire specimen was destroyed. With a 2-hr exposure time at a nominal gas-stream temperature of 1600 F, satisfactory roughness records were usually obtained. The equipment is sufficiently sensitive to measure the corrosion due to diesel oil with a 2-hr exposure at 1600 F.

Briefly then, the place of the surface-analysis method in corrosion testing lies in short-time tests where local temperatures vary according to a known pattern. For long-time tests, where the specimen is maintained at a constant temperature, corrosion must be measured by means of weight-loss measurements. These long-time tests also provide information as to the effect of ash composition on the rate of deposition on blades.

Each short-time test consisted of a 1-hr warm-up period during which diesel oil was burned with a gas temperature of 1200 F followed by a 2-hr test at a nominal temperature of 1600 F with the test fuel. When this latter fuel was a residual oil it was heated and recirculated for several hours before each test to mix thoroughly the ash and any additives which it might contain. At the end of the run the fuel lines and nozzle were flushed out with diesel oil. No measurable corrosive effect was noted due to diesel oil at 1200 F. Specimens were prepared by surface grinding, lapping, and polishing to a 2 microinch RMS finish. Following a corrosion test the specimens were descaled electrolytically in a molten mixture of 60 per cent sodium hydroxide and 40 per cent sodium carbonate. No appreciable change in surface roughness due to this descaling process could be detected. The final surface traverse was made along a band about 1/4 in. from the leading edge. Weight loss also was determined by weighing before and after the test, but these measurements were usually too small to be significant. The readings from the 16 thermocouples were recorded continuously during the test interval by a high-speed multipoint recorder, and were used to determine the temperature profile at the test section. These readings showed the profile to be quite steady, possibly becoming slightly more uniform during the first hour. At intervals during the test, a complete set of readings was taken in order to determine the completeness of combustion, since poor combustion conceivably might affect the corrosion data. In general, the completeness of combustion was 95 per cent on diesel oil and about 90 per cent on residual fuel at the test condition.

A typical isotherm plot for the test section is shown in Fig. 5. A surface-record characteristic of highly corrosive fuels is shown in Fig. 6. Using such surface records, curves of increase in roughness versus temperature can be plotted. Fig. 7 shows such curves for several refractory alloys exposed to the products of combustion of a high-vanadium oil. Fig. 8 shows similar data for chromium platings, chromium having shown some resistance to attack in crucible tests.

Other corrosion investigations have indicated effects due to the presence of carbon (4). The curves in Fig. 9 show the decrease in surface roughness when the same fuel is burned with improved completeness of combustion. This may be due either to the direct effect of carbon on corrosion or to local overheating when carbon deposits burn away.

Surface measurements are essentially two-dimensional in character since the path of the pickup stylus at any time is a straight line and it can sense only the depth and width of a corroded spot. On the other hand, weight loss may be considered to be a three-dimensional function, implying that the surface-roughness data should be raised to the 3/2 power for comparison with weight losses. This is equivalent to assuming that the corroded spot is a hemisphere, and another width dimension is introduced. Such data, further revised with a constant conversion or density factor, are shown in Fig. 10. The specimens from which these data were obtained had been exposed in a furnace, at a very constant temperature, to the combustion products of several fuels. Consequently, the corrosion was very uniform and the correlation with weight loss was easily made. The greatest departure of the converted surface-roughness curves from those for the measured weight loss occurs near the origin; it appears that the proposed 3/2 power conversion is not applicable in this

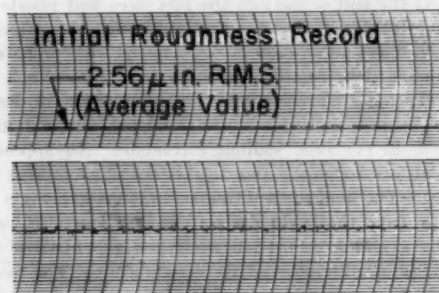
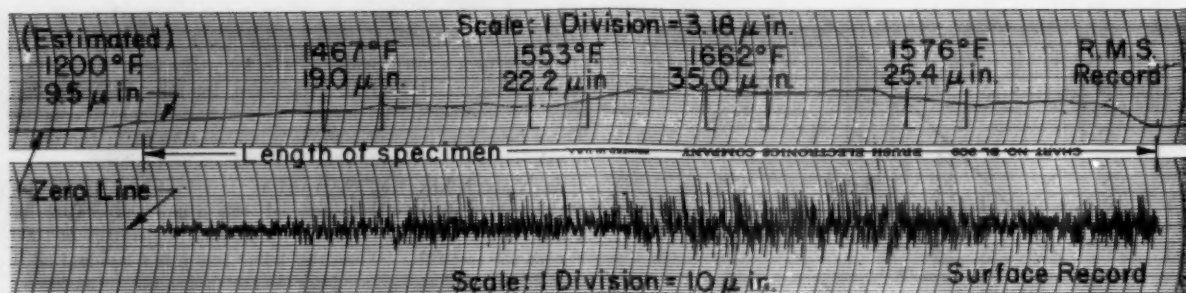


FIG. 6 (Above and left) SURFACE-ROUGHNESS RECORDS BEFORE AND AFTER TESTING WITH HIGH-VANADIUM FUEL

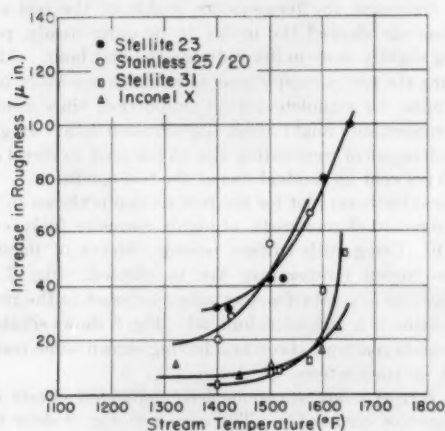


FIG. 7 INCREASE IN ROUGHNESS FOR SEVERAL ALLOYS EXPOSED FOR 2 HR TO COMBUSTION PRODUCTS OF HIGH-VANADIUM FUEL

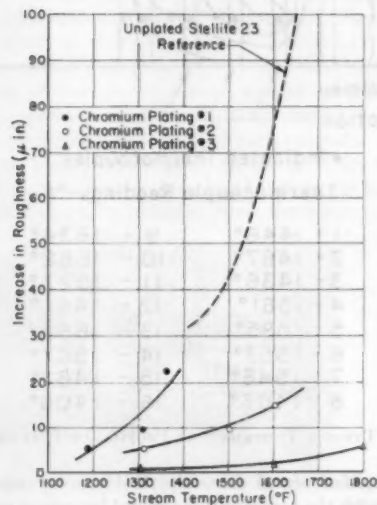


FIG. 8 INCREASE IN ROUGHNESS FOR THREE CHROMIUM-PLATED, STELLITE 23, PIECES, EXPOSED SIMULTANEOUSLY FOR 2 HR TO PRODUCTS OF COMBUSTION OF HIGH-VANADIUM OIL

region, since here it is represented as a parabolic curve rather than a straight line.

ASH-FUSION DETERMINATIONS

It is apparently well-established that catastrophic attack will occur only when an alloy is exposed to the liquid phase of an ash in the combustion products of a corrosive oil (2). And further, that the primary effect of an additive is to raise this ash-fusion temperature. It was, therefore, desirable to have an apparatus for the rapid and accurate determination of fusion temperatures. Such a device should require very small samples, preferably the ash obtained from 100 grams of fuel, and the sample should be visible during the entire determination.

Fig. 11 is a line drawing of the equipment which meets these requirements satisfactorily. The furnace is a modification of a low-temperature design using transmitted light (10). It consists of mating Inconel blocks, each containing a heating coil of

Kanthal A wire set in Norton RA 1162 cement. The furnace is surrounded by B&W 23 insulating brick and covered with a water-cooling cell to protect the microscope. The sample, usually about 20 mg, is placed in a platinum cup located between the heating coils. In the bottom of the cup is a convex dimple containing a thermocouple with leads passing through the bottom of the furnace.

A sircen-arc illuminator provided sufficient intensity for examination up to furnace temperatures of 2000 F, but above this temperature it was necessary to substitute a carbon-arc lamp which permitted observations up to a furnace temperature of 2300 F. With either light source, however, the stray illumination in the microscope barrel was very troublesome. For this reason a stereoscopic microscope was substituted for the monocular

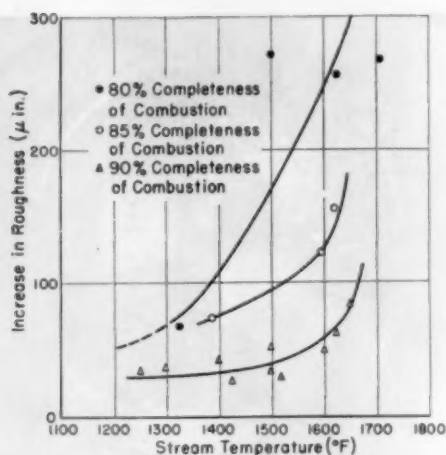


FIG. 9 EFFECT OF COMBUSTION EFFICIENCY ON INCREASE IN ROUGHNESS OF STELLITE 23 AFTER 2 HR OF EXPOSURE TO COMBUSTION PRODUCTS OF HIGH-VANADIUM OIL

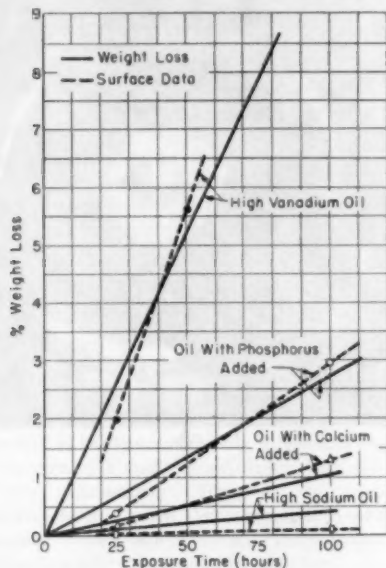


FIG. 10 COMPARISON OF WEIGHT LOSS AND SURFACE-ROUGHNESS MEASUREMENTS FOR 25/20 STAINLESS STEEL AT 1600 F

type illustrated. This microscope, in addition to having a very long working distance, had its illuminating prism located between the objectives but outside the microscope barrel.

The technique of ash preparation consists of igniting and burning out the volatile and tar at as low a temperature as possible after which the carbon residue is heated at 900 F in a muffle furnace until it attains constant weight. It is particularly important that the final ashing temperature should not exceed 1100 F as sodium compounds are apt to volatilize and the sulphates may decompose. To date this method has given excellent and reproducible results.

Four phases of melting have been noted and proposed as standards in evaluating fuel-oil ashes. Photomicrographs taken through the microscope during the actual fusion process are shown in Fig. 12. Initially the ash is granular in form and is lightly packed in the cup. At some relatively low temperature, usually

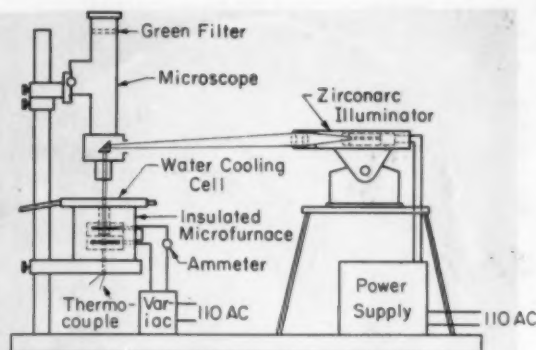


FIG. 11 APPARATUS FOR ASH-FUSION DETERMINATIONS

around 1200 F for high-vanadium oils, or 1400 F for sodium-bearing oils, the ash begins to sinter and pull away from the sides of the cup. If the furnace is cooled at this point, the ash, although still granular, has a crusty texture indicating that some component has fused and holds individual particles together. If heating is continued, the sample reaches a point where small bubbles appear on its surface. The sample is now becoming noticeably darker and tacky. These first two points, arbitrarily defined as the sinter point and first bubble point, are probably significant in determining the deposit-forming characteristics of an ash but in themselves are not directly related to corrosion. Further heating produces an initial melting point where liquid actually can be seen collecting on the bottom of the cup, and eventually the sample melts completely. This last point is the hardest to determine since it is always possible that some solid material remains under the liquid surface. Melted sodium ashes are usually transparent so that solid residue may be seen, but molten vanadium ashes are black and opaque, concealing any solid residue. The initial and final melting points indicate the temperature range where catastrophic corrosion is likely to start. It was noted further that liquid-vanadium compounds effectively wet the metal surfaces and tend to climb out of the cup while sodium compounds usually do not exhibit this tendency. In many tests it is possible to see individual crystals melt and the liquid collect. It is hoped eventually to collect these liquid components as they appear, for microchemical analysis.

After initial calibration runs, ashes from all the fuels tested to date were examined in an attempt to establish the correlation between ash-fusion temperature and corrosivity. Fig. 13 shows this correlation graphically. Each oil is represented by four bars on the lower half of the chart, each bar indicating a phase of melting. Corresponding to these bars, on the upper chart are two bars indicating the relative corrosion of the combustion products when the oils were burned in the high-pressure combustion test rig. The corrosion due to each oil is compared to that of the high-vanadium fuel which is taken as 100 per cent. The corrosion bars were plotted from surface-roughness data taken at 1400 and 1600 F. They show the expected trend of less corrosion for an increase in ash-fusion temperature. If the initial melting point is taken as a corrosion criterion, it will be seen that as long as this temperature is lower than the mean test temperature, 1500 F in the case of these tests, corrosion is severe. But where this temperature is above 2000 F, corrosion is, in general, negligible.

The first ash-fusion results shown at the extreme left of Fig. 13 are for ash collected from a boiler which was fired with the high-vanadium residual oil. The sintering temperature and the other three fusion temperatures for this ash agree very closely with the corresponding temperatures found for the ash formed

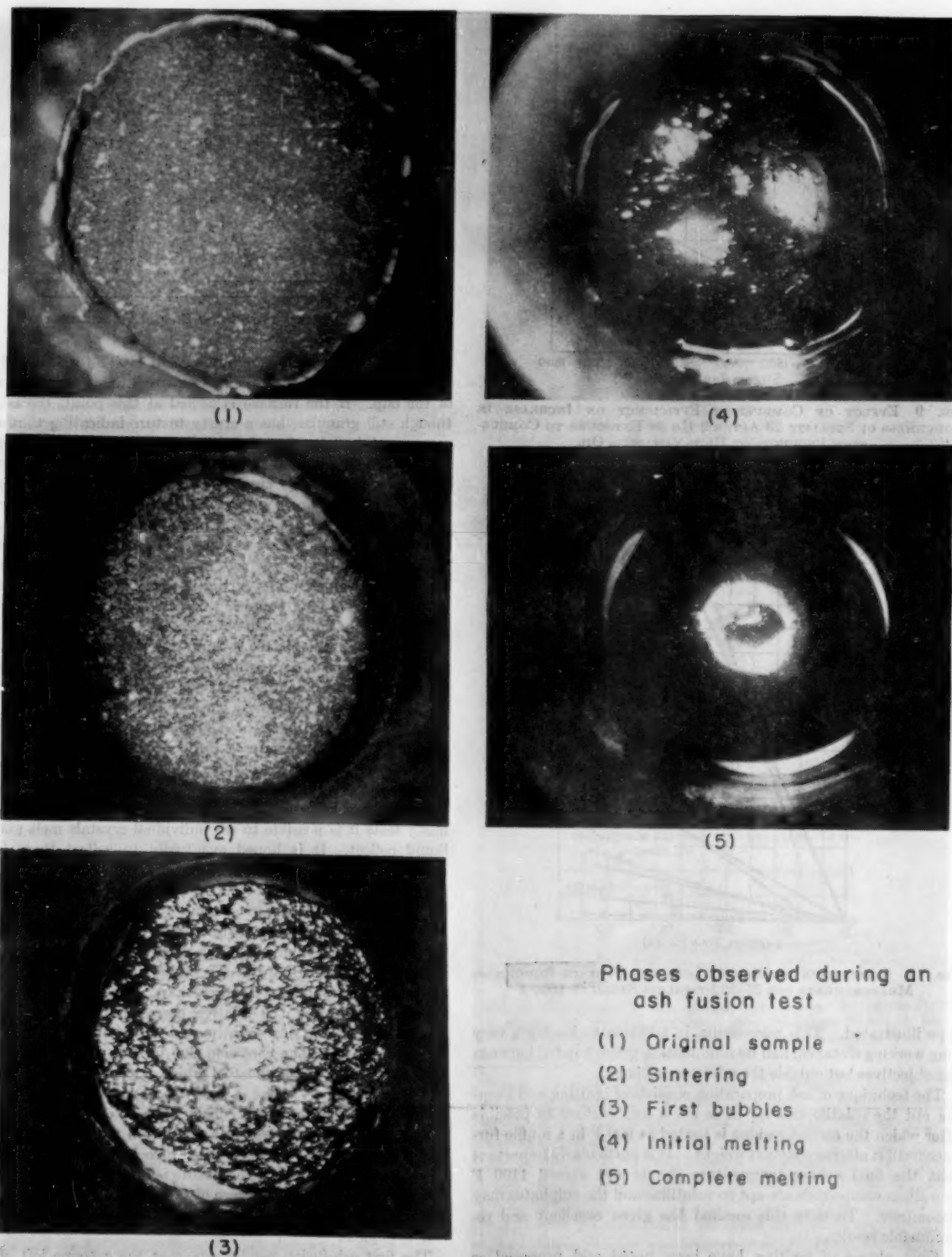


FIG. 12 PHASES OBSERVED DURING AN ASH-FUSION TEST

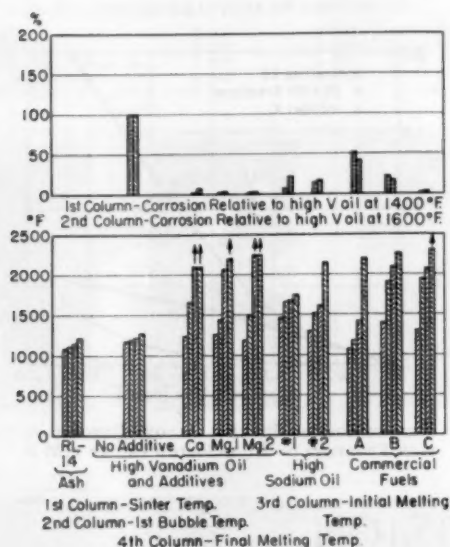


FIG. 13 Ash-Fusion Temperatures and Relative Corrosivity of Several Oils and Oils With Additives

from the high-vanadium fuel by means of the proposed ashing technique. Thus, for this oil at least, the constituents in the laboratory ash appear to be the same as those formed when the oil is burned normally. Similar checks for oils having other ash compositions are to be made. However, it should be noted that the use in corrosion tests of synthetic ashes formed by mixing some of the known constituents together is a questionable procedure.

DISCUSSION OF TEST RESULTS

The investigation of residual-oil corrosion was started when a gas-turbine locomotive was being operated on a number of different railroads and it was necessary to select satisfactory fuels from a large number of sources of supply. Never more than one or two sample drums of fuel from each supply were available and selection had to be made after only a few short tests. The short-time corrosion test and the ash-fusion test met these requirements very satisfactorily. However, the continued use of these tests has demonstrated that considerable additional information bearing on the basic aspects of the corrosion problem can be obtained in a convenient and rapid manner.

Both the similarities and dissimilarities of the corrosive effects due to the presence of vanadium and sodium in residual-oil ash have been mentioned. Since most oils contain both these elements, corrosion with any untreated oil is likely to be a composite effect. To determine the characteristics of the corrosion due to these two ash components a detailed investigation with a high-vanadium and two high-sodium residual fuels was undertaken. Table 2 lists the more important properties and the chemical composition of the ashes for these three oils.

Results With High-Vanadium Oil. Fig. 7 shows the increase in roughness for four typical refractory alloys: 25 Cr-20 Ni stainless steel, Stellite 23, Stellite 31, and Inconel-X, when subjected to a short-time corrosion test burning the high-vanadium fuel. The compositions of these three alloys are given in Table 3. The corresponding results when a Ca and two different Mg compounds were mixed with this fuel before being burned are shown in Fig. 14. The decreased corrosion with any of these additives should be noted. A plausible explanation is supplied by the effect of these additives on the ash-fusion temperatures shown in Fig. 13

TABLE 2 FUEL-OIL SPECIFICATIONS

Property	High vanadium	High sodium-A	High sodium-B
Gravity API	13.7	15.5	14.2
Flash (PM), deg F	208	186	186
Viscosity SS Furol at 122 F	186	162	165
Sediment (extraction), per cent			0.01
Sediment (hot filtration), per cent	0.04	0.14	0.18
Ash total, per cent	0.08	0.20	0.21
Sulphur, per cent	2.51		1.36
Ash analysis ppm based on oil, per cent:			
Al	20	10.6	4
Ca	0	68.5	130
Cu			0.74
Fe	16	23.8	22
Mg	4	18.0	10
Ni	34	14.2	9
S	31	425	
Si	7.2	13.1	16
Na	86	571	600
V	367	17.9	6

TABLE 3 REFRACTORY-ALLOY COMPOSITION

Element	25 Cr-20 Ni	Stellite 23	Stellite 31	Inconel-X
Co	...	65	54	...
Cr	23-26	23-29	23-28	15
Fe	51	2.0	2.0	7.0
Mo
Ni	19-22	1.6	0-12	73
W	...	4-7	0-9	...
Al	0.7
C	0.25	0.35-0.50	0.45-0.60	0.04
Cb	1.0
Mn	2.0	0.5
P	0.04
S	0.03
Si	1.5-3.0	0.4
Ti	2.5

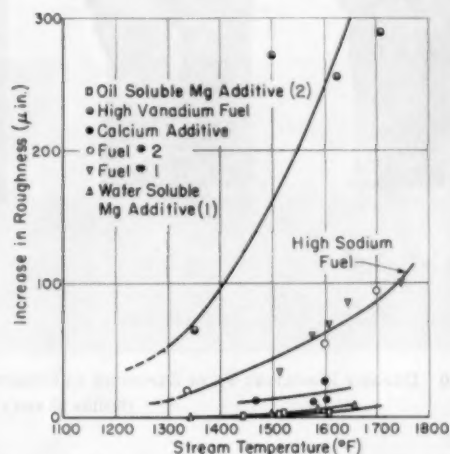


FIG. 14 INCREASE IN ROUGHNESS OF STELLITE 23 EXPOSED FOR 2 HR TO COMBUSTION PRODUCTS OF HIGH-VANADIUM OIL, WITH AND WITHOUT ADDITIVES, AND TWO HIGH-SODIUM OILS

since, with all three, the initial and final melting temperatures of the untreated oil are increased by more than 700 deg F.

When long-time corrosion tests in the atmospheric test rig were run with the same alloys and fuels the foregoing results were confirmed for the Mg additives, as shown in Fig. 15, but the weight loss found after 40 hr with the Ca additive exceeded that found for the untreated oil. The source of this increased corrosion probably lies in the extremely heavy deposit which accumulated on the test specimens when the Ca was added to the fuel as shown in Fig. 16. Sulzer (8) also has observed this condition and offers the following explanation: CaO forms during combustion and collects on the specimen surface where it reacts exothermally with the SO_2 in the gas stream to form $CaSO_4$ (if SO_2 is present in the stream instead, V_2O_5 will catalyze it to SO_3 at the surface). Although the $CaSO_4$ does not melt below 2600 F the heat of formation may raise the temperature suffi-

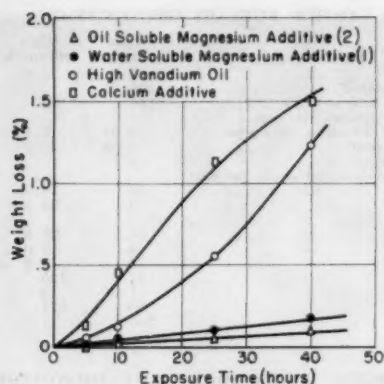


FIG. 15 WEIGHT LOSS OF STELLITE 23 EXPOSED TO COMBUSTION PRODUCTS OF HIGH-VANADIUM OIL WITH AND WITHOUT ADDITIVES AT 1450 F

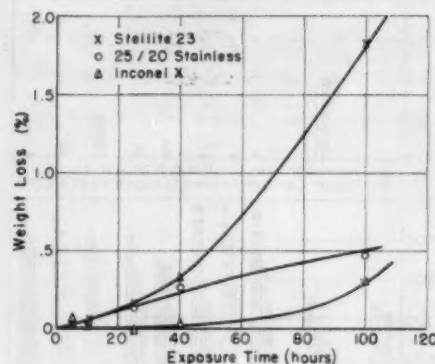


FIG. 17 WEIGHT LOSS OF SEVERAL ALLOYS EXPOSED TO COMBUSTION PRODUCTS OF HIGH-SODIUM OIL AT 1450 F

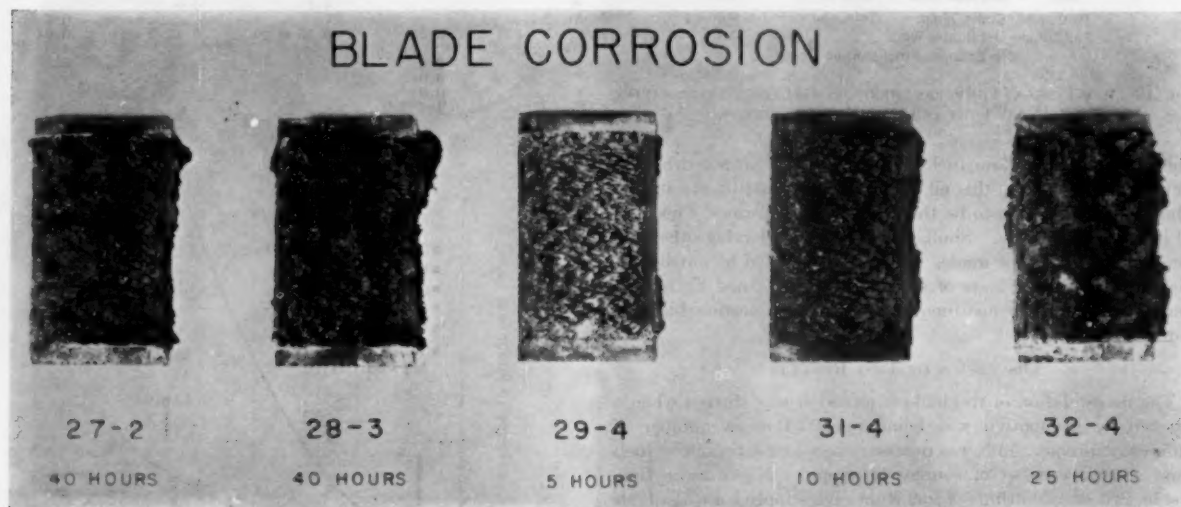


FIG. 16 DEPOSIT RESULTING FROM EXPOSURE TO COMBUSTION PRODUCTS OF HIGH-VANADIUM FUEL CONTAINING CALCIUM ADDITIVE (Stellite 23 test strips; 1400-1500 F; upstream face.)

ciently to cause the surface to become tacky, thus collecting an excessive deposit on the surface of the specimen. Further investigation is needed to confirm this hypothesis.

A number of other refractory materials have been subjected to short-time tests with the untreated high-vanadium oil, but no satisfactory high-temperature alloy has shown superior resistance to this type of corrosion than Inconel-X. The long-time tests and operating experience have corroborated these results in a satisfactory manner.

Results With High Sodium Oil. The increase in surface roughness during a short time corrosion test with the high-sodium oil is shown in Fig. 14. The results lie between those for the untreated high-vanadium oil and those for the latter oil with Ca and Mg additives. The ash-fusion temperatures for the high-sodium oil, as shown in Fig. 13, are found to lie between 1500 and 2100 F. The significance of the initial melting temperature became evident when long-time corrosion tests were run.

The first of these was run at a nominal temperature of 1450 F with the weight losses shown in Fig. 17. While the corrosion of the Stellite 23 occurred more rapidly than was the case with the treated high-vanadium oil, for the 25 Cr-20 Ni and the Inconel-X

the corrosion was not excessive, being less than 0.5 per cent after 100-hr exposure. When the nominal temperature was increased to 1600 F the stainless steel was not affected appreciably, the corrosion of the Stellite was more than doubled, while catastrophic corrosion practically destroyed the Inconel-X as shown in Fig. 18. These results were so unexpected that the test was repeated with almost identical results.

To obtain a better understanding of the corrosive attack with residual fuel having high sodium concentrations in the ash and to explore the possibilities of alleviating this condition with an additive, a comprehensive series of ash-fusion measurements were made. Since Mg additives had proved effective in reducing the corrosion from high vanadium, three Mg compounds were selected as follows:

MgO and $MgSO_4$ added to high-sodium-oil ash.

MgO and magnesium naphthenate added to the fuel oil which was then ashed. Table 4 gives the concentrations of these different additions.

Lack of space prevents the inclusion of a detailed discussion of these ash-fusion results, but from them it was concluded that the addition of a slurry containing MgO in an amount to give an

TABLE 4 CONCENTRATIONS OF ADDITIVES

Mg/Na ratio.....	0.5	1.28	2.0
Gm MgO/gm ash.....	0.237	0.606	0.945
Gm naphthenate/50 gm oil...	0.476	1.22	1.91
Gm MgSO ₄ /gm ash.....	0.462	1.185	1.850
Gm MgO/50 gm oil.....	12.7	31.7	49.6

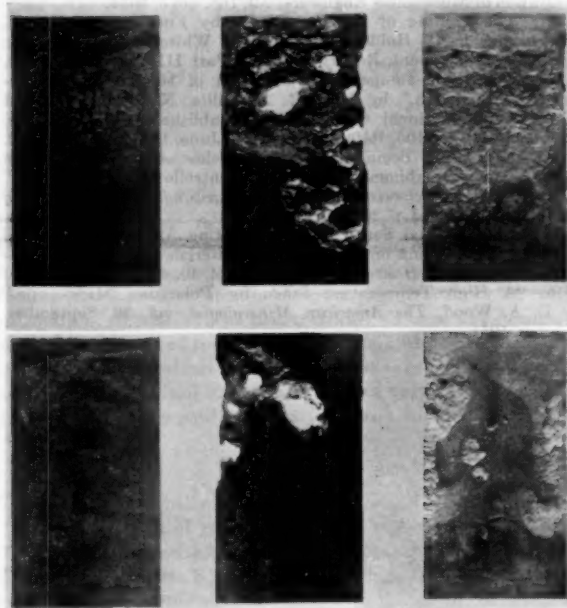


FIG. 18 CORROSION OF SEVERAL ALLOYS RESULTING FROM AN EXPOSURE FOR 50 HR TO COMBUSTION PRODUCTS OF HIGH-SODIUM FUEL AT 1600 F

Mg/Na ratio of 2.0 to the high-sodium oil should produce a reduction in the corrosive attack. The results of 50-hr tests burning this fuel oil without and with the MgO additive are given in Table 5. Comparison of the percentage weight loss given in the last column of this table gives satisfactory confirmation of this conclusion. Even more convincing evidence of the effects of the MgO additive is given in Figs. 18 and 19 which show the specimens from the two tests after descaling. Fig. 20 is a photograph of the test section at the conclusion of the test with the additive. A large deposit has accumulated on the blades, but this was porous and friable and could be easily scraped off. The Ca concentration in this fuel was 130 ppm which probably contributed to the deposit as in the case of the high-vanadium oil to which a Ca addition had been made.

TABLE 5 RESULTS OF TESTS

Specimen no.	Material	Fuel	Per cent weight loss
128.....	25 Cr-20 Ni	High Na	0.7
129.....	Inconel-X	High Na	Total
130.....	Stellite 23	High Na	2.9
131.....	25 Cr-20 Ni	High Na + MgO	0.3
132.....	Inconel-X	High Na + MgO	0.3
133.....	Stellite 23	High Na + MgO	0.3

SUMMARY

It has been the primary purpose of this paper to describe apparatus and test methods which were developed to aid in the investigation of the corrosion associated with the combustion of residual fuel oils in gas turbines. After using these methods for about two years it has been demonstrated that:

1 Short-time tests with corrosion measured by the increase in surface roughness in conjunction with microash-fusion de-

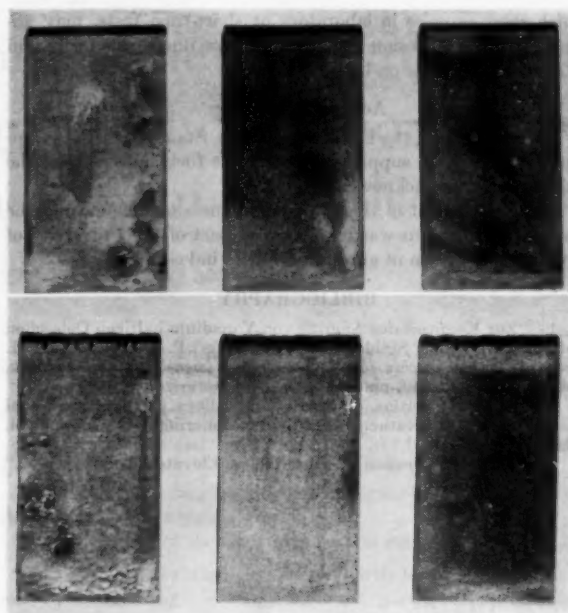


FIG. 19 DESCALED SPECIMENS AFTER 50 HR EXPOSURE TO COMBUSTION PRODUCTS OF HIGH-SODIUM OIL WITH MAGNESIUM-OXIDE ADDITIVE AT 1600 F

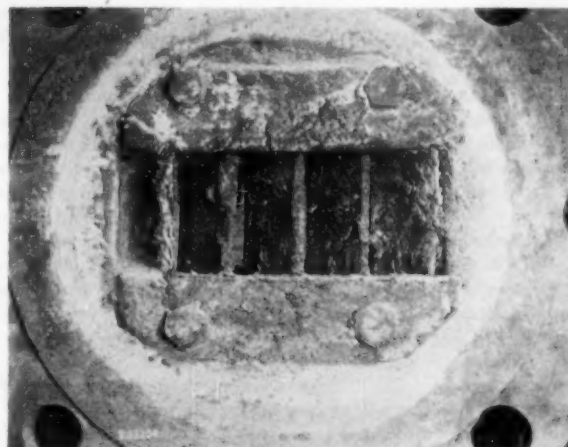


FIG. 20 DEPOSIT FORMED DURING 50 HR RUN ON HIGH-SODIUM OIL WITH MAGNESIUM-OXIDE ADDITIVE; UPSTREAM FACE

terminations are rapid and convenient means of assessing the potential corrosive properties of residual fuel oils. This conclusion has been substantiated by operating experience with a gas-turbine locomotive, and by long-time laboratory tests with combustion rigs operating at atmospheric pressure.

2 Vanadium in the form of V_2O_5 and sodium, probably in the form of Na_2SO_4 , are the most prevalent sources of corrosion with residual fuels. Catastrophic corrosion occurs with the former between 1200 and 1300 F and with the latter between 1600 and 1700 F. Of the more common refractory alloys Inconel-X has about the best resistance to vanadium attack, but has very poor resistance to Na_2SO_4 . MgO in suitable proportions will reduce corrosion due to both these agents.

3 Considerable care and judgment must be used in the selection of fuel additives to reduce corrosion, for many additives,

which show promise in laboratory or short-time tests, may actually increase corrosion in actual operation due to the formation of excessive deposits on blades.

ACKNOWLEDGMENTS

The assistance of the Esso Laboratories, Standard Oil Development Company, in supplying some of the fuels and some of the analytical data is acknowledged.

The development of the surface-roughness techniques and the ash-fusion apparatus was carried on as part of the University of Pittsburgh program of graduate work in industry.

BIBLIOGRAPHY

- 1 "Zur Kenntnis des Angriffs von Vanadium haltigen Oelaschen auf Hitzebeständige Stähle," by P. Schlapfer, P. Amgwerd, and H. Preis, *Schweizer Archiv für Angewandte Wissenschaft und Technik*, vol. 15, Heft 10, 1949, pp. 291-299.
- 2 "Rapid Oxidation of Metals and Alloys in the Presence of MoO_3 ," by G. W. Rathenau and J. L. Meijering, *Metallurgia*, vol. 52, 1950, p. 167.
- 3 "Oil Ash Corrosion of Materials at Elevated Temperatures," by C. T. Evans, Jr., Special Technical Publication no. 108, American Society for Testing Materials, 1950.
- 4 "Residual Fuel-Oil Ash Corrosion," by B. O. Buckland, C. M. Gardiner, and D. G. Sanders, ASME Paper No. 52-A-161.
- 5 "Corrosion Aspects of the Vanadium Problem in Gas Turbines," by S. H. Frederick and T. F. Eden, *Proceedings of The Institution of Mechanical Engineers*, vol. 168, no. 3, 1954.
- 6 "The Fouling of Turbine Blades by Fuel Ash," by J. J. MacFarlane, F. H. Holderness, and F. S. Whitcher, National Gas Turbine Establishment, Report no. R114, Part III, February, 1952.
- 7 "The High Temperature Corrosion of Heat-Resisting Materials by Fuel Ash," by J. D. R. M'Quillin, N. Stephenson, and T. A. Taylor, National Gas Turbine Establishment, Reports nos. R93, R98, R103, R105, R109, R110, R118, June, 1951, to June, 1952.
- 8 "Über die Beeinflussung der Oelaschenablagerungen in Industriellen Gasturbinenanlagen durch Kontrolle der Verbrennung," by P. T. Sulzer, *Schweizer Archiv für Angewandte Wissenschaft und Technik*, Heft 2, 1954, p. 20.
- 9 "Experimental Setup for Investigating Ash Deposition and Corrosion Phenomena on Gas Turbine Materials at High Temperatures," by C. Kind, *Brown Boveri Review*, vol. 40, 1953, p. 196.
- 10 "A High Temperature Stage for Polarizing Microscope," by E. A. Wood, *The American Mineralogist*, vol. 36, September, October, 1951.

The Influence of Some Chemical and Physical Factors on the Formation of Deposits From Residual Fuels

By P. T. SULZER,¹ WINTERTHUR, SWITZERLAND

The oil-ash-deposit problem in industrial gas-turbine plants is analyzed from both physical and chemical standpoints. Theory and experiments on the formation of ash during combustion enable important processes in the build-up of deposits to be understood. These are demonstrated by means of practical examples. Measurements of ash-deposit formation as a function of temperature, pressure, excess air, fuel-ash content, and test duration, are shown graphically and summarized in a formula.

INTRODUCTION

THE primary cause of oil-ash deposits and corrosion during the operation of gas-turbine plants on heavy oil is known to be the organic and inorganic metal-compound content of these fuels, which is often considerable. These form stable metal oxides or chemical compounds with the air in combustion processes using excess air. Furthermore, residual oils contain considerable quantities of sulphur in the form of oil-soluble organic compounds.

The sulphur content is seldom less than 1 per cent of the fuel by weight; the ash content is seldom more than 0.1 per cent. In flue gases the sulphur is in the form of gaseous SO_2 or SO_3 , depending on the temperature. The elements which occur most commonly in oil ash are vanadium, sodium, calcium, iron, nickel, silicon, and aluminum.

A very important part is played by the vanadium and sodium compounds which thus arise. The V_2O_5 in its fluid phase above its relatively low melting point of 650 C acts as a powerful corrosive agent for all metallic alloys and also as a binding agent in building up deposits on the turbine blades (1 to 6).² The sodium constituents, occurring mostly as sulphates, cause very large deposits, but are usually corrosive only in the presence of V_2O_5 .

In spite of the great efforts which have been made in various countries to solve the ash problem, very little is known as yet about the mechanism of deposit formation. Compared with the simpler experimental investigations of ash corrosion, the study of deposit formation presents great difficulties.

In what follows some theoretical observations will be made and then three different kinds of deposit formation will be described with the aid of practical examples. Finally, tests on an experimental rig will be used to explain the formation of deposits quantitatively.

¹ Research Engineer, Sulzer Bros., Ltd.

² Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Gas Turbine Power Division and presented at a joint session of the Gas Turbine Power and Fuels Divisions at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 14, 1954. Paper No. 54–A-171.

THE ORIGIN OF FREE OIL ASH DURING COMBUSTION

The combustion of a single oil droplet in the combustion chamber may be considered to take place in stages as follows (7, 8):

- (a) Evaporation and ignition of the low-boiling-point constituents.
- (b) Chemical decomposition of the higher hydrocarbons by cracking and carbon separation.
- (c) Combustion of the remaining carbon matrix.

The ash constituents in heavy oil are partly in the form of solutions of metallic organic compounds and partly suspensions or emulsions of aqueous solutions of inorganic substances. In the combustion process all the inorganic compounds with relatively low vapor pressure are concentrated in the residual-carbon particle. It is during the complete combustion that the ash particles first begin to form, consisting of stable metallic oxides. The reaction temperatures may be about 1600 C.

One must now briefly distinguish between the following:

(a) Primary reaction of the ash constituents after combustion of all the carbon. The probability of such a reaction is large, since the ash constituents are contained a priori in the oil droplet, and during its combustion they are in intimate contact with one another. All reactions which would form sulphates are absent due to the high temperatures.

(b) Secondary reactions, which occur in the gas stream between the flame zone and the place of deposition, usually in the temperature range 1400–700 C. In this zone the chances of reaction between the solid ash particles are extremely small, but between solid particles (according to their size) and SO_2 or SO_3 rather greater, and between vapor-phase ash constituents and SO_2 or SO_3 they are greatest.

Chemical equilibrium probably does not exist on account of the generally high flow velocities and the continual temperature change. Furthermore, the occurrence of turbulence and contact between the gases and the oxidized metal surfaces (catalytic effect, etc.) often have a decisive effect on the development of the reaction.

(c) Tertiary reactions at the deposition surface. Here the reactions between the deposited ash and the metal are of interest as well as those between the deposits and SO_2 and SO_3 (possibly CO_2 if very little SO_2 is present).

As it is not possible to consider all these reactions, the mechanism of a few important processes only will be more fully discussed. It is convenient for this purpose to divide heavy oils arbitrarily into three classes according to their ash composition:

Fuels With High Vanadium Content. These are rich in vanadium—containing porphyrin complexes. During the final stage of the combustion process vanadium pentoxide V_2O_5 is formed which is stable in an oxidizing atmosphere below 1200 C. On account of its considerable vapor pressure ($>2.5 \times 10^{-4}$ atm at 1000 C) all the V_2O_5 would be evaporated at high temperature under the pressures obtaining in most gas turbines, in the absence of other

ash constituents (9). Fouling by condensation, according to the local pressure and temperature conditions, is then to be expected.

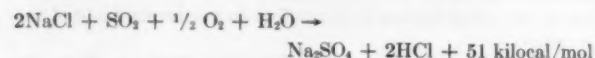
These oils in the first group, however, generally contain besides vanadium considerable quantities of iron, nickel, sodium, calcium, etc. Depending on the ash composition, V_2O_5 will form vanadates or other compounds with the constituents mentioned partly during the primary reaction. The vapor pressures of all the complex compounds which can arise are without exception much lower for a given temperature than that of pure V_2O_5 . It is seen therefore that the greater part of this ash reaches the turbine as rather large particles from the combustion of a single oil droplet.

It will be shown later, by means of an example, that condensation of V_2O_5 can always be confirmed.

It is well known that the deposition of vanadium-containing ash is appreciably less when incomplete combustion is maintained. We have investigated this relationship thoroughly, both theoretically and experimentally, and have drawn the following conclusions:

The oxidation of the organic vanadium compounds in the fuel to V_2O_5 takes place in the final phase (in the residual-carbon matrix) via the lower oxide stages V_2O_3 and V_2O_4 or their mixed phases. In contrast to V_2O_5 , both V_2O_3 and V_2O_4 have a very high melting point and a very low vapor pressure. If now the air supply or the atomization in the combustion chamber is so adjusted that the carbon particle arising from the disintegration of an oil droplet no longer burns completely, then oxidation of the lower-vanadium oxides to V_2O_5 does not occur. The particles, consisting of the lower-vanadium oxides mixed with further ash constituents and carbon, pass through the turbine as a result of their low adhesive properties as a harmless dust and cause neither corrosion nor deposits. The efficacy of this process is therefore particularly good with fuels having a high vanadium content. Larger quantities of sodium compounds act unfavorably, as these have a low softening point both under oxidizing and reducing conditions.

Fuels With a High Alkali Content. Some fuels contain considerable quantities of sodium, up to 0.2 per cent, mostly as an emulsion of an aqueous solution of NaCl, or also as a suspension of microcrystalline particles arising from the emulsion. The probability that NaCl is evaporated during combustion is much greater than for V_2O_5 for two reasons; NaCl has a much higher vapor pressure, at 1000 C about 4×10^{-2} atm. V_2O_5 arises only in the lower-temperature zones, after complete combustion of all the carbon, while NaCl is already present at very high temperatures. NaCl is not stable in the presence of SO_2 and SO_3 . It reacts with SO_3 exothermally according to the following equation



A secondary reaction between gaseous NaCl and SO_2 according to the foregoing equation is therefore probable. This may be indirectly proved by the fact that with sufficiently high sulphur content NaCl deposits never occur. Na_2SO_4 has a much lower vapor pressure than NaCl ($< 1.3 \times 10^{-3}$ atm at 1000 C). The deposits due to Na_2SO_4 are formed therefore by a precipitation of the finest particles or drops arising from the foregoing reaction.

Fig. 1 shows the fouling of a cascade of turbine blades in the outlet ducts of a combustion chamber after a 24-hr combustion test with a fuel containing 0.2 per cent NaCl and 2.5 per cent sulphur (10). The operating pressure was 1.5 atm abs, the gas temperature at the cascade 650 C. Over 95 per cent of the deposits consist of Na_2SO_4 . Although the melting point of Na_2SO_4 lies well above the gas temperature at about 884 C, the layers are

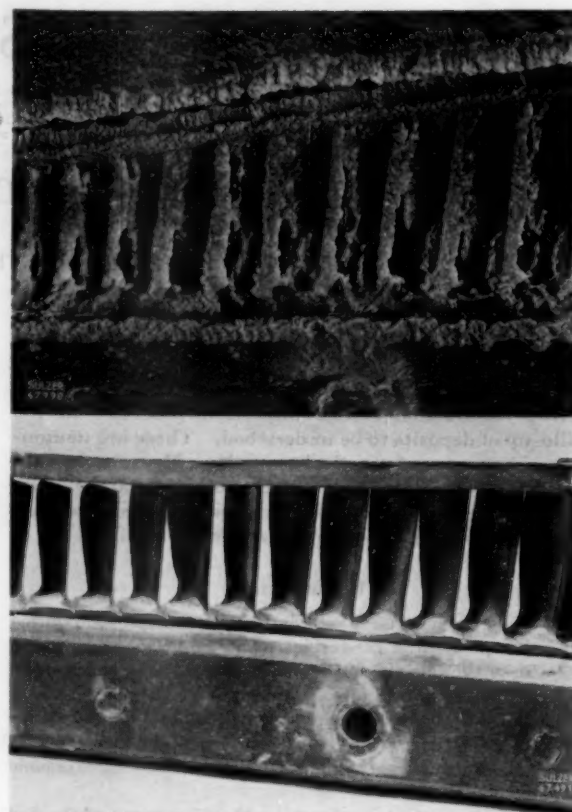


FIG. 1 Na_2SO_4 DEPOSITS ON TURBINE BLADES
(a, After 24 hr residual oil test in experimental rig. b, Clean turbine blade.)

sintered hard or molten. An explanation for this is the considerable heat the particles are subjected to in the exothermic sulphate reaction (see equation).

Fuels With High-Silicon-Content Ash. The ash of certain coal-tar oils consists mainly of aluminum silicates; i.e., of fine suspensions of clay-like substances. In addition there are generally iron and possibly calcium. Appreciable deposits are not to be expected, as the resulting particles have a very high softening point of over 1100 C.

In Fig. 2 are shown curves for a typical fuel of each of the three groups just defined giving quantity of ash deposited against the corresponding temperature. A description of the experiments and their execution is given further on.

Curve 1 is typical of the behavior of oil ash with a high vanadium content. The dependence on temperature is very marked, as the particles receive no additional heat from exothermic secondary or tertiary reactions. Data on the composition of the ashes are given in Table 1.

Curve 2 shows the fouling due to the ash of a residual oil with high sodium content. The flatter shape of this curve is due to the additional heat evolved by the particles reacting with SO_2 and SO_3 .

TABLE 1 COMPOSITION OF ASHES

Per cent	Fuel		
	1	2	3
Ash content.....	0.05	0.2	0.18
Sulphur.....	1.9	2.5	0.5
V_2O_5	0.025
Na_2O	0.01	0.09	..
$SiO_2 + Al_2O_3$	0.14

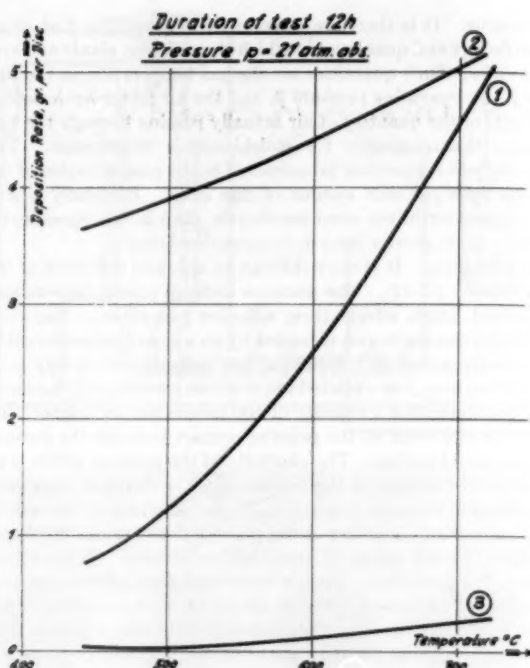


FIG. 2 DEPOSITION CURVES OF THREE TYPICAL RESIDUAL OILS

Curve 1 = oil with high vanadium content
 Curve 2 = oil with high sodium content
 Curve 3 = oil with high silicon content

Curve 3 shows that the deposits from a fuel in group 3 are insignificant.

PREVENTION OF OIL-ASH DEPOSITS BY FUEL ADDITIVES

All studies and investigations in connection with the oil-ash problem are made with the object of finding a suitable process for trouble-free burning of residual oils. Among other methods, the mixing of additives appears to be promising.

It is endeavored to so change the ash artificially, that neither deposits nor corrosion occur; i.e., the additive should raise the melting point of the ash and combine with the chemically most dangerous constituents.

Exhaustive practical tests have been carried out with the semiclosed-cycle 20,000-kw gas-turbine plant of the North East Swiss Power Company (N.O.K.) at Weinfelden in Switzerland, on the combustion of residual oils with and without additives (11). Very favorable results were obtained by adding a suspension consisting mostly of aluminum silicates. The total running time using treated fuels was 2200 hr. During all this time it was not found necessary to clean the turbine; no reduction in performance due to fouling of the turbine could be detected.

Fig. 3 shows the first three stages of the stator blading in the high-pressure turbine after a 50-hr test run without additives. The fuel contained 0.05 per cent ash, consisting of 0.03 per cent V_2O_5 and 0.01 per cent Na_2O , together with a further 1.9 per cent of sulphur. From these results operation with residual oil appeared to be out of the question for the time being.

Fig. 4 shows the same stages after running for 1200 hr with a similar fuel using additives under even more severe operating conditions. The pressure at the turbine inlet was 25 atm abs and the gas temperature 680 C.

The deposits visible in Fig. 3 consist mainly of ash with dissolved corrosion products. The consistency is typical of a de-

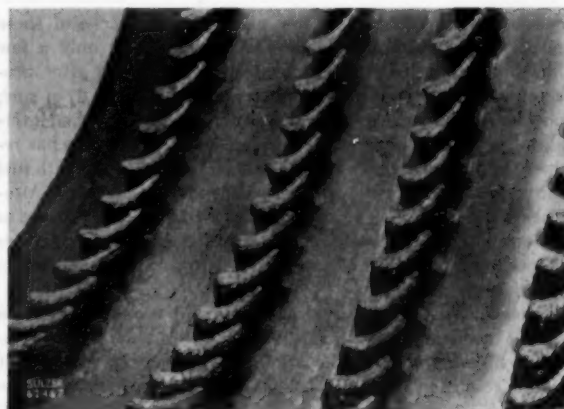


FIG. 3 FOULING OF FIRST THREE STAGES OF H-P TURBINE AFTER 50-Hr TEST WITH RESIDUAL OIL. MAXIMUM BLADE TEMPERATURE, 650 C; OPERATING PRESSURE, 18 ATM ABS

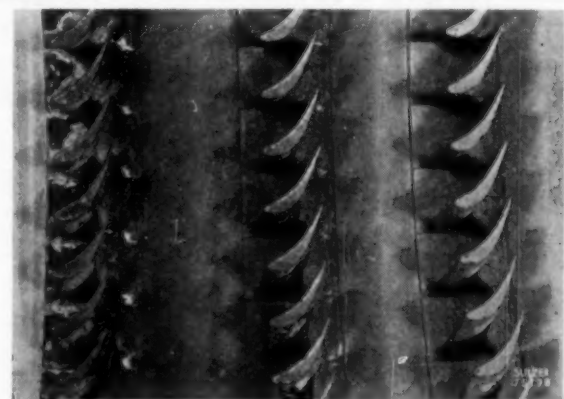


FIG. 4 FIRST THREE STAGES OF H-P TURBINE AFTER 1200 Hr OPERATION WITH RESIDUAL OIL PLUS ADDITIVE. MAXIMUM BLADE TEMPERATURE, 680 C; OPERATING PRESSURE, 25 ATM ABS

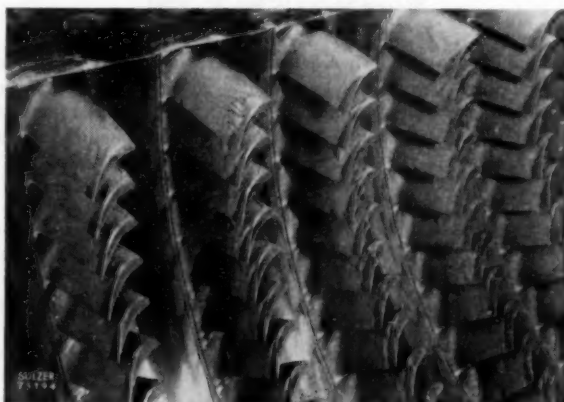


FIG. 5 REAR SURFACE OF FINAL STAGES OF H-P TURBINE BLADES AFTER 1200 Hr OPERATION

posit formation in which each completely burned oil droplet gives a single plastic ash particle.

Fig. 5 shows the final stages of the high-pressure turbine. The pressure is about 14 atm abs, the temperature about 600 C.

Fine-powder deposits are clearly visible on the backs of these blades. The maximum thickness is attained after only a few hours' operation and does not increase further even after protracted running. In the $\times 250$ photomicrograph (Fig. 6) of such a powder specimen, glass-clear globules can be distinguished interspersed with needle-like crystals. Each of these globules has resulted from the combustion of a single oil droplet mixed with fuel additive, and consists of over 90 per cent molten silicates. Altogether about 60 to 80 per cent of the ash is chemically combined in all these globular ash particles. The melting point of the particles is about 1300 C; i.e., high enough to prevent any adhesion in the turbines.

From analyses of various dust fractions it is seen that 20-30 per cent of the available V_2O_5 is still evaporated before it is combined by the additive in the primary reaction. In the present case the conditions in the final stages of the high-pressure turbine appear to be such that partial condensation of the free V_2O_5 occurs. Specially evolved methods of microanalysis show that the needle-like crystals clearly visible in Fig. 6 consist of pure V_2O_5 . It is assumed that V_2O_5 condenses on the nonplastic dust particles while still in the gas stream and that these thereby acquire a certain adhesive property.

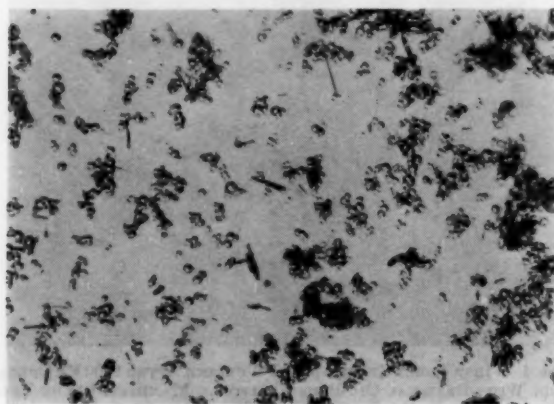


FIG. 6 PHOTOMICROGRAPH OF DEPOSITS ON REAR SURFACE OF BLADES SHOWN IN FIG. 5; $\times 250$

Although the powder-like deposits do not cause any trouble in service, they are nevertheless slightly corrosive owing to the considerable vanadium content. It has now been found possible to eliminate this subsidiary effect by means of a small change in the additive.

EXPERIMENTAL TECHNIQUE

Nomenclature

- T = absolute temperature, deg C
- p = absolute pressure, atm abs
- λ = air factor = excess-air ratio + 1
- α = ash content
- z = duration of test
- C_1, C_2 = constants
- G = weight of deposits

Reports on Operation. Reports on operating experience with residual-oil fuels in industrial gas turbines are often contradictory, even when the fuels are similar with regard to ash content and composition. As already mentioned, the completeness of combustion can have considerable effects in certain cases.

Incidence of Ash Fouling. The engineer and designer must be able to estimate how prone a given gas-turbine system will be to

ash fouling. It is therefore convenient to determine first of all those factors and quantities in which the different plants and systems differ. Such quantities are the gas temperature in the turbine T , the operating pressure p , and the air factor λ ; λ defines the ratio of the quantity of air actually passing through the turbine to that necessary for stoichiometric combustion. The quantity p/λ is therefore proportional to the concentration of the ash particles per unit volume of flue gases. Evidently p/λ is usually greater in the semiclosed-cycle than in the closed-cycle which is again greater than in an open-cycle plant.

Flow Velocity. It is more difficult to appraise the effect of the flow velocity (7, 12). The particles undergo plastic deformation on impact which affects their adhesive properties. Also each particle in coming to rest is heated by an amount proportional to its velocity squared. Although this quantity of energy converted into heat, when related to the whole particle, only amounts to a few degrees for a velocity of 100 meters per sec, higher temperatures can occur at the point of contact between the particle and the metal surface. The plasticity of the particle, which is an exponential function of the temperature, is therefore increased; the adhesion strength is greater. If the plasticity of the ash at the temperature obtaining at the place of deposition is itself considerable (fuels in group 1), then the flow velocity will have little effect. The particles in this case have very good adhesive properties at very low velocity or even simply by contact alone. Much more important than the flow velocity (this falls within a relatively narrow range for the various industrial-plant systems) are the effects of local turbulence conditions on the fouling process. These determine to a very great extent what proportion of the available particles will come into contact with the surface.

During the tests described later, the flow velocity could be held constant within about 20 per cent. Special tests for further investigating these effects quantitatively are being projected.

Other necessary quantities are the ash content α and the duration of test z .

In the accompanying diagrams the test results have been summarized as the quantity of deposit formation in an experimental plant in terms of temperature T , concentration p/λ , ash content α , and the test duration z .

ARRANGEMENT OF TEST RIG

The high-pressure combustion test rig used in the present tests is shown in Fig. 7. Two piston compressors with a total output of 1200 kg/hr supply the compressed air for combustion. The maximum pressure is 25 atm abs. This air first passes through the recuperator 6 and is then led to the combustion chamber 1 where it is divided into two main streams, primary and secondary air. Small quantities are tapped off for cooling the burner, inspection windows, etc. The atomizer with spill return enables a maximum fuel flow of 25 kg/hr to be injected. Ignition is effected by insertion of a glowing metal coil.

After leaving the combustion chamber the flue gases pass a cascade consisting of 7 slotted disks, Fig. 8, which serve as deposition surfaces. In order to provide a large temperature drop along the cascade, the disks are connected by water-cooled pipes. It is possible to obtain a temperature drop of over 400 deg C depending on the temperature adjustment at the first stage. The amount of ash deposited is determined by weighing the separate disks.

A single test at some definite pressure is sufficient to determine the deposit formation over the whole temperature range which is of interest. After leaving the cascade the gases enter the cooler 4. The pressure is adjusted by means of the valve 5. Thence the flue gases exhaust through the heat exchanger 6.

In Fig. 9, the combustion chamber appears in the foreground, with the cooler and air heater to the rear. All the measuring instruments and control apparatus are mounted on the wall. The

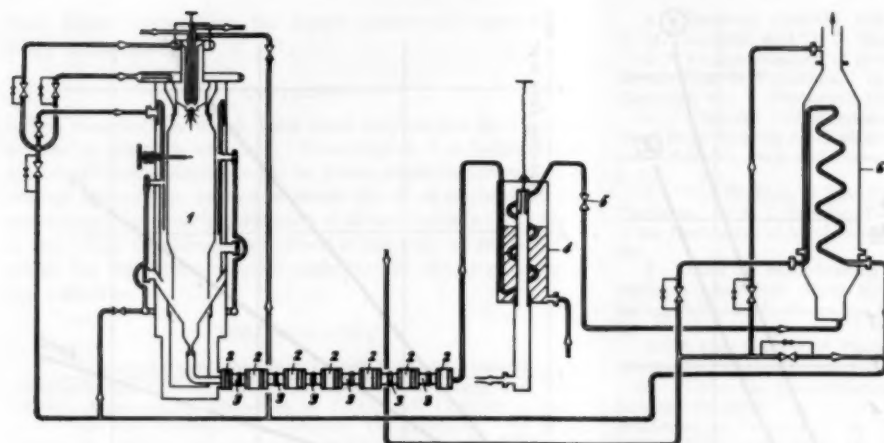


FIG. 7 DIAGRAM OF EXPERIMENTAL HIGH-PRESSURE TEST RIG

- 1 Combustion chamber
- 2 Slotted deposition disks
- 3 Water cooling
- 4 Cooler
- 5 Valve
- 6 Air preheater

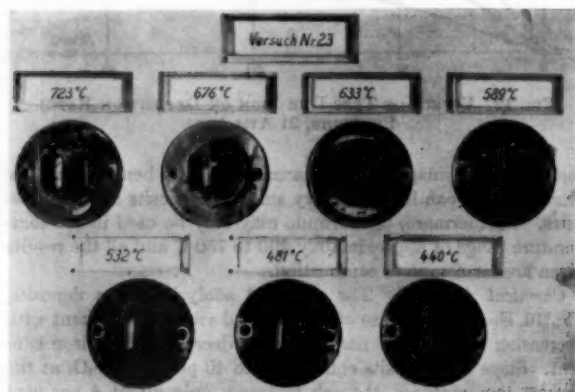


FIG. 8 DEPOSITION DISKS AFTER 3-Hr TEST SHOWING DEPOSITION AT VARIOUS TEMPERATURES

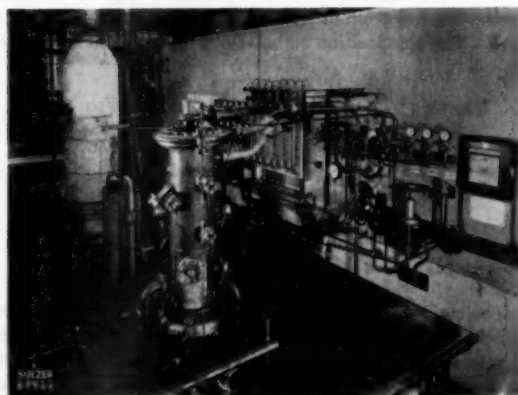


FIG. 9 GENERAL VIEW OF HIGH-PRESSURE COMBUSTION TEST RIG—EXCLUDING COMPRESSORS

deposition cascade can be inspected by removing panels in the floor.

Unfortunately, it was not possible to use the same fuel for all the series of tests; at least all three oils belong to group 1 and are similar both in ash content and composition.

RESULTS

Analysis of the family of curves in Fig. 10 shows at once that

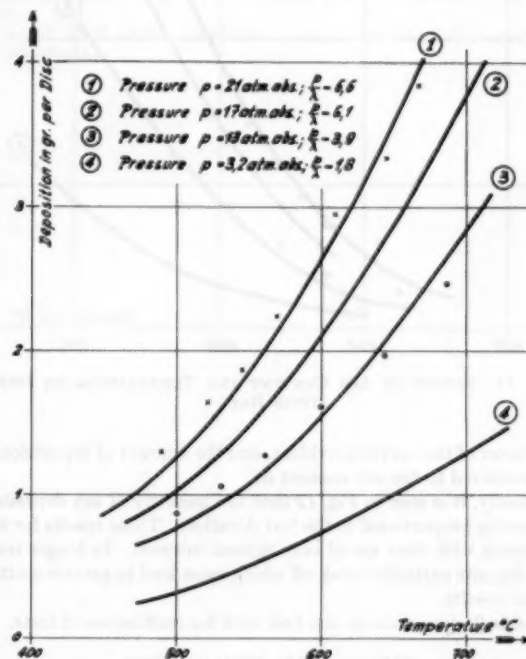


FIG. 10 EFFECT OF TEMPERATURE AND CONCENTRATION P/λ ON DEPOSITION RATE

deposition increases as an exponential function of the temperature and also that it is proportional to the concentration p/λ . Each curve was determined by burning a constant amount of fuel. The curves 1, 2, and 3 are based on a 12-hr test, curve 4 on a 23-hr test. The exponential characteristic is explained by the variation of particle plasticity with temperature, which follows the law $e^{-(A/T)+B}$ with good approximation. Here A and B denote constants, T the absolute temperature. In the higher temperature range the deposits have a greater tendency to break away together with the metal surface, which gives considerable scatter in the results.

The effect of ash content is found from the family of curves in Fig. 11. Curve 1 corresponds to a test with undiluted residual fuel, while curves 2, 3, and 4 show the results for the same fuel diluted with an ash-free distillate. Here, also, the exponential

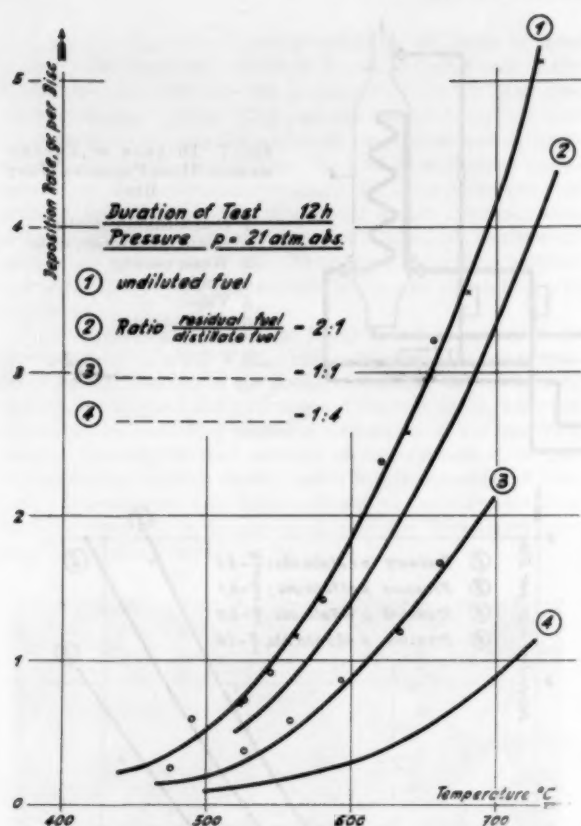


FIG. 11 EFFECT OF ASH CONTENT AND TEMPERATURE ON DEPOSITION RATE

character of the curves is evident, and the amount of deposition is proportional to the ash content α .

Finally, it is seen in Fig. 12 that the quantity of ash deposited is directly proportional to the test duration. These results for the variation with time are of only limited interest. In longer tests the deposits partially break off which must lead to greater scatter in the results.

Table 2 gives data on the fuel used for each series of tests.

TABLE 2 FUEL USED IN TESTS

Per cent	Fuel		
	Fig. 9	Fig. 10	Fig. 11
Ash content.....	0.05	0.05	0.04
Sulphur.....	1.9	1.8	1.9
V ₂ O ₅	0.025	0.03	0.018
Na ₂ O.....	0.009	0.005	0.012
Fe ₂ O ₃	0.002	0.001	0.002

The results of the three series of tests can be summarized by the following formula, as may be easily proved

$$G = C_1 \alpha x \frac{p}{\lambda} e^{-C_2/T}$$

The constants C_1 and C_2 are determined for each fuel by means of a single test. All the curves shown can be calculated with good approximation by this means.

It is evident that the range of application for this formula is very limited, especially for fuels in which the ash contains large quantities of earth-alkali metals, where considerable deviations have been found. This effect is explained by the marked dependence on pressure of the sulphate formation. For instance,

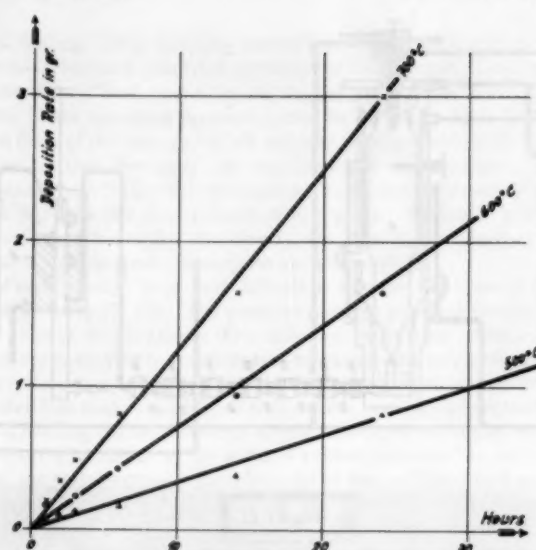


FIG. 12 EFFECT OF TEST DURATION ON DEPOSITION RATE; PRESSURE, 21 ATM ABS

the presence of magnesium is harmless or even beneficial at low pressures, but can lead to heavy sulphate deposits at high pressures. Furthermore, the formula may only be used in the temperature range of approximately 400 to 750 C and all the results given are for complete combustion.

Chemical Analysis. The chemical analysis of the deposits, Fig. 10, show an increase of sulphur and vanadium content with decreasing temperature and a marked decrease in the iron content. Since the deposits contain up to 40 per cent Fe₂O₃ at the highest temperatures, the ash however, only about 4 per cent (Table 2), it is obviously a matter of corrosion products coming from the surface. If these values are converted on the basis of a 4 per cent Fe₂O₃ content, then a continuous, slightly increasing alkali and sulphur content with decreasing temperature is found, Table 3. The composition is, however, independent of the pressure.

TABLE 3 ANALYSIS OF ASH DEPOSITS (FIG. 10)

Per cent	Temperature	
	700 C	500 C
V ₂ O ₅	48	40
(Na ₂ O + K ₂ O).....	9	12
SO ₂	21	29
Fe ₂ O ₃	4	4
SiO ₂	8	6
Remainder.....	10	9

* Converted to 4 per cent.

This behavior can be explained as follows: From the foregoing it is probable that the sodium compounds contained in the fuel evaporate and form Na₂SO₄ particles. The greater part of the V₂O₅ together with the remainder of the ash constituents (Fe₂O₃, SiO₂, NiO, CaO, Al₂O₃) form a plastic particle after complete combustion of the oil particle. Thus the deposition of Na₂SO₄ particles is superimposed on the development of deposits of relatively large particles with a pronounced plasticity variation with temperature, whereas the plasticity of the former is apparently much less temperature-dependent, Fig. 2. Thus in the lower temperature range relatively more sodium sulphate is deposited. This effect has been observed much more clearly in deposition processes with ash having a higher Na₂O:V₂O₅ ratio. The curves found consist then of two simple exponential curves superimposed. For this reason the points in Fig. 10 also deviate to-

ward higher values from the simple exponential curve at the lower temperatures.

CONCLUSION

The observations which have been summarized here do not pretend to scientific accuracy. Nevertheless, it is believed that they enable some conclusions to be drawn about the interaction of various phenomena, as well as about the effect of the most important quantities in the formation of oil-ash deposits. The study of the rather complicated processes is not only of academic interest, but has given valuable assistance in the preparation of fuel additives.

BIBLIOGRAPHY

- 1 "Untersuchungen über den Angriff V-haltiger Oelaschen auf hitzebeständige Stähle," by P. Schlöpfer, P. Angwerd and Preis, *Schweiz. Archiv angew. Wissensch. u. Techn.*, vol. 15, 1949, p. 291.
- 2 "Oil Ash Corrosion at Elevated Temperatures," by C. T. Evans, Jr., Symposium, American Society for Testing Materials, 1950.
- 3 "Einige Probleme der warmfesten, hitzebeständigen Stähle von Standpunkt des Verbrauchers," by W. Stauffer, *Schweiz. Archiv angew. Wissensch. u. Techn.*, vol. 17, 1951, p. 353.
- 4 "Residual Fuel-Oil Ash Corrosion," by B. O. Buckland, C. M. Gardiner, and D. G. Sanders, ASME Paper No. 52-A-161.
- 5 "An Investigation of Accelerated Oxidation of Heat Resistant Metals Due to Vanadium," by F. C. Monkman and N. J. Grant, *Corrosion*, vol. 9, December, 1953, pp. 460-466.
- 6 "Über die V_2O_5 -Korrosion sunderbeständiger Werkstoffe und deren Beeinflussung durch einzelne Legierungselemente," by E. Fitzer and J. Schwab, *Berg- und Hüttenmännische Monatshefte*, vol. 98, 1953, p. 1.
- 7 "The Problem of Fuel-Oil Ash Deposition in Open-Cycle Gas Turbines," by A. T. Bowden, P. Draper, and H. Rowling, *Proceedings of the Institution of Mechanical Engineers*, series A, vol. 167, 1953, p. 291.
- 8 "Über die Beeinflussung der Oelaschenablagerungen in industriellen Gasturbinen durch Kontrolle der Verbrennung," by P. T. Sulzer, *Schweiz. Archiv angew. Wissensch. u. Techn.*, vol. 20, 1954, p. 33.
- 9 P. Lloyd and R. P. Probert, *Proceedings of The Institution of Mechanical Engineers*, vol. 163, 1950, p. 206.
- 10 "Über die Beeinflussung von Oelaschenablagerungen durch Brennstoffzusätze," by P. T. Sulzer, *Schweiz. Archiv. für angew. Wissensch. u. Techn.*, vol. 18, 1952, p. 379.
- 11 "Brennstoffzusätze zur Verhinderung von Oelaschenablagerungen in Gasturbinen," by P. T. Sulzer, *Schweiz. Bauzeitung*, vol. 72, 1954, p. 1.
- 12 M. C. G. Smith, unpublished work at the National Gasturbine Establishment.

Compressibility Deviations for Polar Gases

By N. A. HALL¹ AND W. E. IBELE,² MINNEAPOLIS, MINN.

The departures of compressibility factors of polar gases from generalized compressibility charts have been noted by previous studies of gas compressibility as reaching such magnitudes as to reduce seriously the utility of such charts in describing the behavior of polar gases. Employing an "extended law of corresponding states," this paper shows that polar-gas compressibility departures correlate with the "reduced dipole moment," a parameter which is closely related to the molecular dipole moment of the particular polar gas. Charts are presented which permit the calculation of compressibility factors for polar gases by applying a polarity correction to the compressibility as given by a standard chart.

INTRODUCTION

A PROBLEM frequently encountered in engineering analysis concerns the prediction of pressure-volume-temperature behavior of a gas under the manifold conditions occurring in various industrial processes. More often than not the data required lie in a range marked by the total absence of direct experimental P - V - T data. The problem promises to be a continuing one as the number of industrially significant gases increases and the need for data extends to regions not yet considered by experimental investigations. Generalized charts, constructed by averaging the experimental compressibility data for several gases, represent the best practical means available for meeting this need, the most familiar examples being those of Dodge (1),³ Su (2), Obert (3), and Haugen and Watson (4). Experience has shown that such charts achieve moderate success when developed according to the law

$$Z = f(P_r, T_r) \dots \dots \dots [1]$$

where $Z = PV/RT$, the compressibility factor and the right side of Equation [1] is a universal function in reduced pressure and temperature. The reduced properties are defined by $P_r = P/P_c$ and $T_r = T/T_c$ where P_c and T_c are the critical values. Increased accuracy may be achieved at the price of decreased generality by developing and using the charts with a particular family of gases as has been done for hydrocarbons (2). When charts were derived as an average of data for several dissimilar gases, the result was not in rigorous agreement with all the data on any one of the gases. These discrepancies resulted from the fact that the theory of corresponding states, as postulated, was only an approximation, and in the case of polar gases and "light gases" not an entirely satisfactory one. Various empirical methods (5, 6, 7) have been proposed to resolve this difficulty but even the most

recent semiempirical study (8) notes deviations as high as 7 per cent for certain gases.

THEORETICAL APPROACH

Concurrent with a renewed interest in empirical developments and modifications of the law of corresponding state has been an approach on a theoretical basis. The principle of corresponding states was formulated in such a way that no reference was made to any particular equation of state. The conditions to be met for substances to conform to the principle were clearly set forth (9) and the following equivalent mathematical statement derived by an exact theoretical treatment results

$$F(V^*, T^*, P^*) = 0 \dots \dots \dots [2]$$

where V^* , T^* , and P^* are obtained from the usual macroscopic quantities pressure P , volume V , and temperature T , by rendering them dimensionless through the use of two molecular constants which refer to the distance (r_0) at which the generalized intermolecular potential has a minimum energy ϵ_0 . Thus $V^* = V/Nr_0^3$, $T^* = \kappa T/\epsilon_0$, $P^* = Pr_0^3/\epsilon_0$, N and κ being Avogadro's number and Stefan-Boltzmann's constant, respectively. Failure of the light gases, hydrogen, deuterium, helium, and to a lesser degree neon to conform to Equation [2], led to the introduction (10) of a light-gas parameter to account for light-gas deviations. Gases possessing permanent electric or dipole moments have been treated successfully (11) by means of an intermolecular potential (12) that involves a parameter related directly to the dipole moment μ of the molecule. Again, gases having nonpolar molecules of cylindrical symmetry have been considered (13) by introducing a "shape" parameter based on the maximum internuclear distance of the molecule.

These activities have led to proposed extensions of the statistical law of corresponding states (11, 14) including as additional factors the dimensionless quantities analogous to the statistical mechanics parameters found to be important in considering light, polar, and nonsymmetrical gases. This study seeks to verify macroscopically the proposed extension for the group of polar gases.

VALIDITY OF PROPOSED EXTENSION

Since the compressibility factor represents one of the most useful properties indicating gas imperfections and is also one of the most accurately and extensively measured quantities associated with gas behavior, recourse was made to this property in testing the validity of the proposed extension; thus

$$Z = f(T_r, \rho_r', \mu_r, l_r, \Lambda_r, \dots, \dots, \dots) \dots \dots \dots [3]$$

ρ_r' representing Su's (2) ideal reduced density defined by $\rho RT_c/P_c$; μ_r , l_r , and Λ_r representing parameters important for polar, nonspherical, nonpolar, and light gases, respectively. The ideal reduced density was chosen in order to avoid the introduction of the critical density, a quantity more subject to error than the critical temperature and pressure. When the interaction of permanent dipoles exerts the dominant influence on gas behavior, Equation [3] reduces to

$$Z = f(T_r, \rho_r', \mu_r) \dots \dots \dots [4]$$

The statistical treatment of polar gases (12) employs a dimen-

¹ Professor of Mechanical Engineering, Heat Power Division, University of Minnesota. Mem. ASME.

² Assistant Professor of Mechanical Engineering, Heat Power Division, University of Minnesota. Assoc. Mem. ASME.

³ Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division and presented at the Annual Meeting, November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 26, 1954. Paper No. 54-A-140.

sionless parameter constructed of microscopic quantities, $\mu^2/\epsilon_0 r_0^3$, which may be converted to macroscopic ones by using the results of Lennard-Jones (15) that $\epsilon_0 \propto \kappa T_c$ and $V_c/N \propto r_0^3$; thus

$$\mu^2/\epsilon_0 r_0^3 \sim \mu^2/(\kappa T_c)(V_c/N) = \mu_r \dots \dots \dots [5]$$

The success of statistical mechanics in representing observed P - V - T data and transport properties by means of intermolecular potentials suggested as a possible standard, that compressibility factor as would be obtained through a generalized intermolecular potential. Such a potential, by its parameters could then be related to the effects that lead to departures from the conventional principle of corresponding states. A search of the literature failed to disclose a potential of such generality. Even if one had been found, or formulated, the standard compressibility predicted would have been subject to two limitations: (a) The compressibility prediction would occur through some equation of state constituting a restricted interpretation of the principle of corresponding states which properly formulated makes no reference to any particular equation of state. (b) Even should a particular equation be chosen as representing the standard, the virial form, for example, the attendant mathematical difficulties have found only the second and third virial coefficients of the series computed to date; thus to select one of the most successful equations of state would, nevertheless, circumscribe the range of application.

RARE GASES SELECTED AS STANDARD

To avoid accepting an equation as representing the principle of corresponding states, the group of rare gases, argon, krypton, and xenon, was selected as the standard. While such a selection was arbitrary, certain advantages accrued from such a choice. It has been shown (16) that these gases form a group which conforms most closely to Equation [1]; consequently, deviations from this model by other gases are equivalent to obtaining departures from the conventional two-parameter statements of the principle. Another gain was realized in terms of the extent and accuracy of the data available (17, 18, 19, 20, 21). The accuracy of the experimental data was of critical importance since in some cases the departures were of a magnitude to be masked by the experimental error had the latter been excessive. The justification for the selection of the model rests therefore on rather pragmatic grounds.

The compressibility of the standard was taken as

$$Z_{STD} = f(T/T_c, \rho RT_c/P_c) \dots \dots \dots [6]$$

where $\rho RT_c/P_c$ represents the "ideal" reduced density. This formulation was chosen to avoid the introduction of plotting errors by minimizing the number of cross plots necessary before representing the data in the final form for comparison. Only one construction chart of Z versus ρ_r at various reduced temperatures was required in order to construct the standard chart of reduced isometrics shown in Fig. 1. In order to preserve the accuracy of the data treated, the plotting scale was selected to realize a plotting uncertainty of not more than two parts per thousand. The vapor-pressure curve shown was calculated from the saturated-vapor properties of argon (22). The critical constants used to place the data in reduced form were those given by (23, 24, 25).

In order that the proposed extension might rest on as broad a base as possible, a comprehensive search of the literature was undertaken. A group of 20 gases was composed from the entire body of gases whose compressibilities have been measured. The elimination of only slightly polar gases, those of unknown dipole moment or insufficient P - V - T data reduced the number to 13. Table 1 summarizes the polar gases employed in the study and the

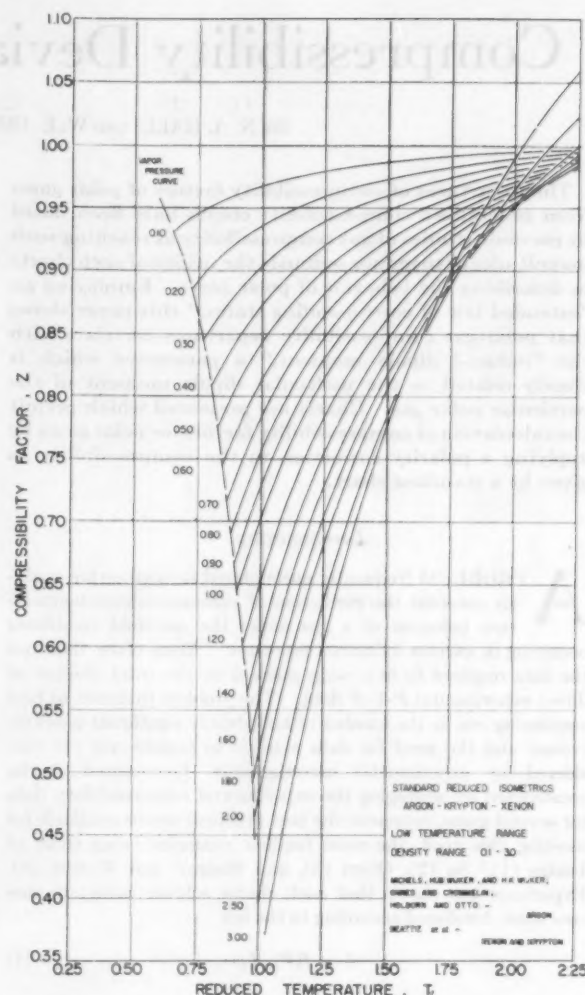


Fig. 1

sources from which the pertinent experimental data indicated were obtained. Clearly, not all the P - V - T data listed are of equal accuracy. The group of freons was studied originally primarily for the purpose of acquiring equation of state and thermodynamic data for the purpose of engineering design and application, activities somewhat less demanding of highly accurate data. Conversely, ammonia and water have been studied extensively, and their behavior over reasonable ranges of pressure and temperature are known to be accurate and consistent. The compressibility data for methyl fluoride represents a recent contribution to the literature from the Leiden group, long renowned for its precision studies of pressure-volume-temperature behavior of gases. The remainder of the gases considered would appear to rank between the two bounds described with respect to accuracy and consistency of the data.

CONSTRUCTION CHARTS

Following the same procedure used in connection with the standard charts, a construction chart in reduced co-ordinates, Z versus ρ_r (ideal reduced density), was made for each of the gases. From these charts, plots were made of compressibility factor versus reduced temperature for reduced isometrics at intervals of

TABLE 1 SUMMARY OF POLAR GASES

Gas	Source of P-V-T data ^a	Critical constants			Sources	Saturated vapor data	Molecular dipole moments	
		Pressure, int. atm	Temperature, deg K	Density, gm/cc			$\mu \times 10^{18}$ esu	Source ^a
Ammonia, NH ₃	26, 27, 28	111.5	405.6	0.235	46	54	1.45	61
Freon 12, CCl ₂ F ₂	29, 30	39.56	384.7	0.556	47	55	0.51	62
Freon 21, CHCl ₂ F.....	31	51.0	451.6	0.522	48	56	1.29	62
Freon 22, CHF ₂ Cl.....	31	48.48	369.6	0.525	4849	57	1.35	62
Ethyl alcohol, C ₂ H ₅ OH.....	32	63.1	516.3	0.276	46	32	1.69	62
Ethyl ether, (C ₂ H ₅) ₂ O.....	33, 34, 35	35.5	467.0	0.262	46	35	1.13	63
Freon 11, CFCl ₃	31	43.2	471.1	0.554	47	58	0.45	62
Hydrogen sulphide, H ₂ S.....	36	58.87	373.6	0.349	36	59	1.02	62
Methyl alcohol, CH ₃ OH.....	37, 38	78.7	513.2	0.272	46	60	1.66	64
Methyl chloride, CH ₃ Cl.....	39, 40	65.89	416.1	0.37	40	40	1.86	64
Methyl fluoride, CH ₃ F.....	41	62.0	318.1	0.318	50	..	1.81	64
Propene, C ₃ H ₆	42, 43	45.47	365.0	0.227	43, 51, 52	42	0.35	62
Water, H ₂ O.....	44, 45	218.2	647.3	0.318	53	53	1.94	64

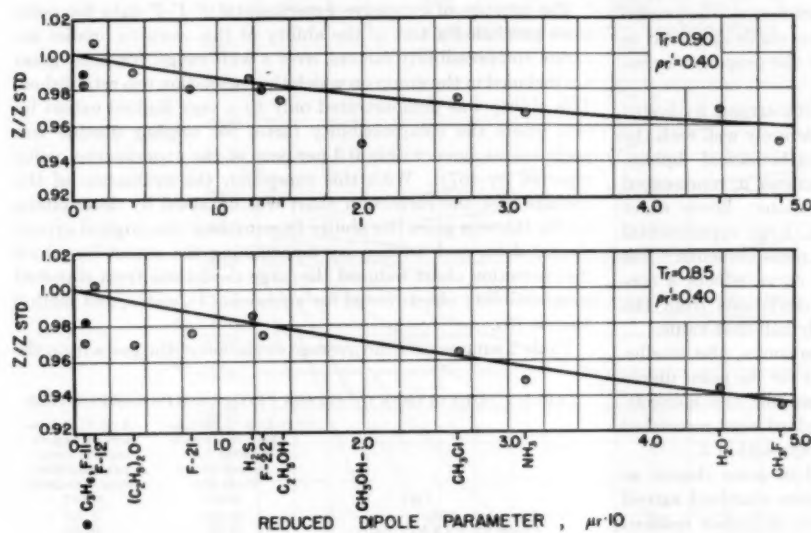
^a See Bibliography.

FIG. 2

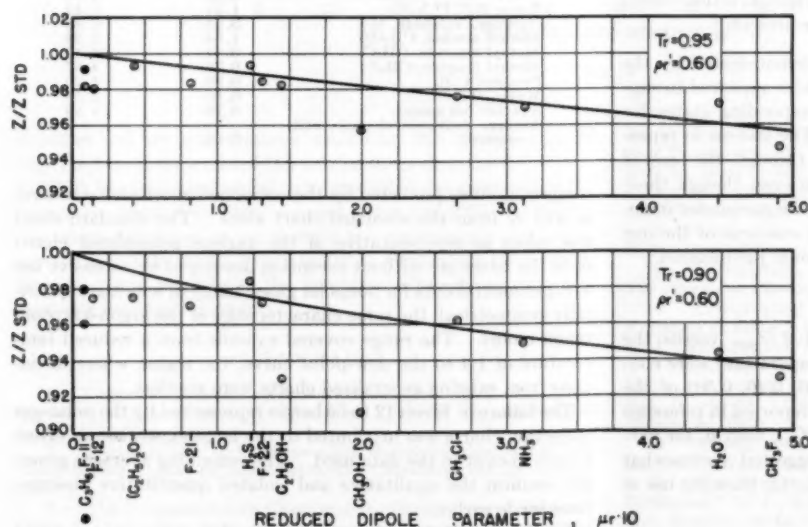


FIG. 3

one tenth from 0.1 to 1.0 and two-tenth intervals from 1.0 to 2.0. The results were similar to Fig. 1 of the model and are given by reference (65).

successfully reproduced the experimental data.

2 The group of freons on the whole conforms surprisingly well to the average curve, particularly when the nature of the experi-

Using Equation [4] as representing the group of thirteen polar gases, the ratio Z/Z_{STD} was formed at identical reduced-density and reduced temperature values for the gases considered and the standard. Since the three gases composing the model obey very closely the two-parameter law of corresponding states, the standard compressibility was taken as defined by the reduced temperature and ideal reduced density; hence

$$Z/Z_{STD} = f(T_r, \rho_r, \mu_r)/f(T_r, \rho_r', 0) \quad \dots [7]$$

To correlate the factor Z/Z_{STD} with the reduced dipole parameter, defined by Equation [5], plots were made with μ_r as abscissa and Z/Z_{STD} as ordinate. This was accomplished at the same density intervals appearing in Fig. 1 and at reduced temperature increments of 0.05 from the dew-point curve to a reduced temperature of 1.20 and at 0.10 increments to 1.40. Lack of data prevented extension to higher reduced temperatures and densities. The validity of the correlation is shown in part by Figs. 2 and 3, the complete results reported by (65).

PERTINENT OBSERVATIONS

Bearing in mind the experimental errors present and the tendency of inherently diverse species to deviate from the average behavior despite the mutual sharing, in various degrees, of a common property, the following observations were made:

1 Propene, with smallest value of reduced dipole parameter, did not correlate well with the other polar gases. An unsaturated hydrocarbon, propene might better fit with gases having cylindrical symmetry, an inference supported by the fact that the Benedict-Webb-Rubin equation (68) usually used for nonpolar hydrocarbons, suc-

mental P - V - T data was recalled. Freons 21 and 22 exhibited, throughout the range of density and temperature covered by experimental compressibility data, as satisfactory agreement as might be wished. Correlation of Freons 11 and 12 was satisfactory except at high densities and low temperature where rather large deviations were expected due to known errors in the compressibility data.

3 Hydrogen-sulphide deviations, which were as large as 3 per cent relative to the standard, always correlate to within 1 per cent of the average curve.

4 For ammonia and methyl chloride the values of Z/Z_{STD} always agreed within $1/2$ per cent with the correlation curve; more often than not the best curve fit passed exactly through the Z/Z_{STD} values for these gases. When it was recalled that the experimental compressibility factors for these gases deviated by more than 7 per cent from standard chart prediction, the precise agreement of these gases with the correlation curve appeared to constitute strong evidence of the validity of the proposed extension of the principle of corresponding states.

5 Methyl alcohol and ethyl alcohol with strong hydrogen bonding tendencies (66) correlated only moderately well with the other polar gases. The idealization of mathematical dipoles, from which evolved the reduced dipole parameter μ_r , represented only an approximation to the real interactions; hence exact agreement was not expected. In addition, large experimental errors are known to exist in ethyl-alcohol measurements. For both substances, however, the correlation curve effects a correction in the proper direction reducing deviations from the standard chart as large as 7 per cent to nearly half that value.

6 Despite strong hydrogen bonding tendencies, the smaller size of the water molecule probably accounts for the point dipole model offering a better representation than for the alcohols. Departures as large as 7 per cent from standard were correlated within 1 per cent in almost all cases and always within 2.

7 Methyl-fluoride departures correlated to same degree as water; deviations as large as 9 per cent from standard agreed within 1 per cent with the curve fit except at higher reduced temperatures where agreement was to within 2 per cent.

8 Ethyl-ether deviations from the correlation seldom exceeded 1 per cent, the corrections being made in the right direction except in the region where experimental errors were present.

In retrospect, the degree of correlation demonstrated by the various gases with the reduced dipole parameter appeared to support the extension of the principle of corresponding states for polar gases represented by Equation [4]. The success in reproducing experimental compressibility data through the use of generalized parameters would be acceptable even though these parameters were entirely empirical. When the parameter introduced has a demonstrated physical basis as was true of the one employed here the extension receives additional justification.

CORRECTION CHARTS

For engineering use, correction charts of Z/Z_{STD} versus the ideal reduced density at various reduced temperatures were constructed for selected values (0.10, 0.20, 0.30, 0.40, 0.50) of the reduced dipole parameter held constant and reported in reference (65). The linear variation of the ratio Z/Z_{STD} with μ_r for various reduced densities and temperatures suggested a somewhat simpler method of obtaining the correction factor than the use of five separate charts; thus

$$Z/Z_{STD} = 1 + \phi(T_r, \rho_r')\mu_r \dots \dots \dots [8]$$

was chosen to represent the behavior of the correction factor, the function $\phi(T_r, \rho_r')$ determined by rearranging Equation [8] to give

$$\phi(T_r, \rho_r') = \frac{Z/Z_{STD} - 1}{\mu_r}$$

With selected values of the reduced temperature and ideal reduced density held constant, a value of $\phi(T_r, \rho_r')$ was obtained for values of μ_r . At the same reduced temperature, the procedure was repeated, but with a different value of the ideal reduced density. This was continued until $\phi(T_r, \rho_r')$ had been evaluated for the initial reduced temperature held constant, at values of the density, ρ_r' , from 0.10 to 2.0. Following a similar process for other values of the reduced temperature allowed Fig. 4 to be plotted. The compressibility for a polar gas is given then by using the reduced co-ordinates of the state to determine Z_{STD} from Fig. 1 and the same co-ordinates with μ_r to determine the value of $\phi(T_r, \rho_r')$, Z may then be determined by Equation [8].

The absence of extensive experimental P - V - T data for polar gases precluded a test of the ability of this work to predict accurate compressibility factors, over a wide range, for polar gases not included in the group on which the correlation was established. This ability was demonstrated only to a very limited extent by (65) where the compressibility factor for sulphur dioxide was predicted to agree within 0.3 per cent of the experimental value reported by (67). With this exception, the evaluation of the usefulness of the correction chart was obtained by determining for the thirteen gases the ability to reproduce the original experimental data, and further, by determining the extent by which the correction chart reduced the large deviations from standard compressibility charts noted for ammonia (1), water, and methyl fluoride (8).

Table 2 summarizes the average deviation of the gases from the

TABLE 2 DEVIATION OF GASES FROM COMPRESSIBILITIES

Gas	Avg % deviation of experimental compressibilities from this work	Avg % deviation of experimental compressibilities from standard chart
Ammonia, NH_3	0.22	2.53
Freon 12, CCl_2F_2	0.86	0.83
Freon 21, $CHCl_2F$	0.53	1.27
Freon 22, CHF_2Cl	0.36	1.46
Ethyl alcohol, C_2H_5OH	1.24	1.60
Ethyl ether, $(C_2H_5)_2O$	0.56	0.86
Freon 11, $CFCl_3$	1.31	1.44
Hydrogen sulphide, H_2S	0.44	0.88
Methyl alcohol, CH_3OH	1.64	3.28
Methyl chloride, CH_3Cl	0.14	2.10
Methyl fluoride, CH_3F	0.76	4.67
Propene, C_3H_6	0.90	1.20
Water, H_2O	0.86	2.56
Average (all gases).....	0.76	1.90
Average (without alcohols and propene).....	0.60	—

compressibilities predicted by the correction chart and standard, as well as from the standard chart alone. The standard chart was taken as representative of the various generalized charts since the latter are without exception developed by extensive use of experimental data for nonpolar gases, many of which are spherically symmetrical, the same characteristics of the argon-krypton-xenon model. The range covered extends from a reduced temperature of 1.4 to the dew-point curve, the region where deviations from existing generalized charts were greatest.

The failure of Freon 12 to be better represented by the polar-gas correction charts was attributed to the large experimental errors known to exist in the data used. The remaining averages generally confirm the qualitative and isolated quantitative observations made earlier.

Further evaluation of the correction charts was accomplished by comparing the predicted compressibility factors against those predicted by the methods of statistical mechanics (11) for ammonia. The results appear in Table 3. The increasing deviation between experimental and statistical values noted at high

TABLE 3 COMPRESSIBILITY FACTORS FOR $\mu_r = 0.312$

$T_r = 0.868$			$T_r = 0.947$				$T_r = 1.026$			
Ideal reduced density			Ideal reduced density				Ideal reduced density			
0.20	0.40	0.50	0.20	0.40	0.60	0.80	0.20	0.40	0.60	0.80
0.8827	0.7654	0.7067	0.9081	0.8164	0.7249	0.6334	0.9258	0.8517	0.7777	0.7038
(S)	(S)	(S)	(S)	(S)	(S)	(S)	(S)	(S)	(S)	(S)
0.8958	0.7973	0.7500	0.9170	0.8398	0.7670	0.6988	0.9329	0.8709	0.8120	0.7563
(E)	(E)	(E)	(E)	(E)	(E)	(E)	(E)	(E)	(E)	(E)
0.8998	0.8021	0.7570	0.9195	0.8411	0.7685	0.6996	0.9346	0.8710	0.8112	0.7545
(I)	(I)	(I)	(I)	(I)	(I)	(I)	(I)	(I)	(I)	(I)

NOTES:

(S), Statistical Mechanics, Bird, Hirschfelder, Curtiss.

(E), Experimental values for ammonia, data of Beattie and Lawrence.

(I), This work

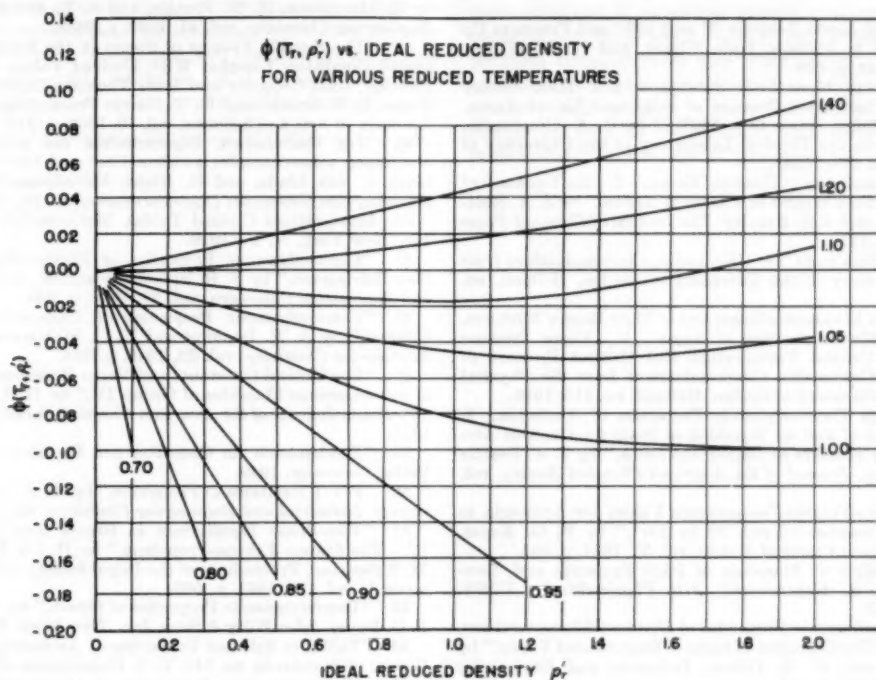


Fig. 4

densities and low temperatures result from the allowable use of only the first three terms of the infinite series form for Z .

While lack of experimental data precluded the extension of the correlation beyond an ideal reduced density of 2.00, the correction chart developed to supplement the accurate standard compressibility chart should prove useful in eliminating large deviations of polar gases from the compressibility charts currently available for engineering use.

BIBLIOGRAPHY

- 1 "Chemical Engineering Thermodynamics," by B. F. Dodge, McGraw-Hill Book Company, Inc., New York, N. Y., and London, England, 1944, p. 160.
- 2 "Modified Law of Corresponding States for Real Gases," by Gong-Jen Su, *Industrial and Engineering Chemistry*, vol. 38, 1946, p. 803.
- 3 "The Compressibility Chart and the Ideal Reduced Volume," by E. F. Obert, *Industrial and Engineering Chemistry*, vol. 40, 1948, p. 2185.
- 4 "Chemical Process Principles," by O. A. Haugen and K. W. Watson, John Wiley & Sons, Inc., New York, N. Y., vol. 2, 1943, p. 489.
- 5 "Activity Coefficients of Gases," by R. H. Newton, *Industrial and Engineering Chemistry*, vol. 27, 1935, p. 302.
- 6 "The Critical Constants and Compressibility of Ethane and a

Study of the Generalized Relations for the Compressibility of Gaseous Hydrocarbons," by Gong-Jen Su, Doctorate thesis, Massachusetts Institute of Technology, Cambridge, Mass., June, 1937.

7 "P-V-T Relations of Gases," by H. P. Meissner and R. Seferian, *Chemical Engineering Progress*, vol. 47, 1951, p. 579.

8 "Generalized $p_r T_r$ Properties of Gases," by L. C. Nelson and E. F. Obert, *Trans. ASME*, vol. 76, 1954, pp. 1057-1066.

9 "Corresponding States for Perfect Liquids," by K. S. Pitzer, *The Journal of Chemical Physics*, vol. 7, 1939, p. 583.

10 "Contribution to the Quantum-Mechanical Theory of the Equation of State. Determination of the Law of Force of Helium," by J. DeBoer and A. Michels, *Physica*, vol. 5, 1938, p. 945.

11 "Survey of the Equation of State and Transport Properties of Gases and Liquids," by R. B. Bird, J. O. Hirschfelder, and C. F. Curtiss, University of Wisconsin, CM-758, Nord-9937 Task Wis-1-c, November, 1952.

12 "Second Virial Coefficient of Polar Gases," by W. H. Stockmayer, *The Journal of Chemical Physics*, vol. 9, 1941, p. 398.

13 "The Second Virial Coefficient of a Gas of Non-Spherical Molecules," by J. Corner, *Proceedings of the Royal Society of London, England, series A*, vol. 192, 1948, p. 275.

14 "States of Matter," by J. A. Beattie and W. H. Stockmayer, edited by H. S. Taylor and Glasstone, D. Van Nostrand Company, Inc., New York, N. Y., third edition, vol. 2, 1951, p. 344.

15 "The Equation of State and Critical Phenomena," by J. E. Lennard-Jones, *Physics*, vol. 4, 1937, p. 941.

16 "Thermodynamics, an Advanced Treatment for Chemists and Physicists," by E. A. Guggenheim, North-Holland Publishing

Company, Amsterdam Interscience Publishers, Inc., New York, N. Y., second edition, 1950.

17 "Isotherms of Monatomic Gases and of Their Binary Mixtures VII. Isotherms of Argon Between $+20^{\circ}$ and -150° C," by H. Kamerlingh Onnes and C. A. Crommelin, Communications from the Physical Laboratory at the University of Leiden, Holland, no. 118b, 1909-1911.

18 "Über die Isothermen einiger Gase bis 400° and ihre Bedeutung für das Gasthermometer," by L. Holborn and J. Otto, *Zeitschrift für Physik*, vol. 23, 1924, p. 77.

19 "Über die Isothermen, von Helium, Stickstoff und Argon unterhalb 0° ," by L. Holborn and J. Otto, *Zeitschrift für Physik*, vol. 30, 1924, p. 320.

20 "Über die Isothermen einiger Gase zwischen $+400^{\circ}$ und -183° ," by L. Holborn and J. Otto, *Zeitschrift für Physik*, vol. 33, 1925, p. 1.

21 "Isotherms of Argon Between 0° and 150° and Pressures Up to 2900 Atm.," by A. Michels, Hub. Wijker, and H. K. Wijker, *Physica*, vol. 15, 1949, p. 629.

22 "Isothermals of Monatomic Substances and Their Binary Mixtures. XV. The Vapor Pressure of Solid and Liquid Argon, From the Critical Point Down to -206° C," by C. A. Crommelin, Communications from the Physical Laboratory of the University of Leiden, Holland, no. 138c, 1931.

23 "The Compressibility of Gaseous Xenon. I. An Equation of State for Xenon and the Weight of a Liter of Xenon," by J. A. Beattie, R. J. Barriault, and J. S. Brierley, *The Journal of Chemical Physics*, vol. 19, 1951, p. 1219.

24 J. J. Meihuizen and C. A. Crommelin, Communications from the Physical Laboratory of the University of Leiden, Holland, no. 245c, 1937.

25 "Isotherms of Monatomic Gases and of Their Binary Mixtures. IV. Remarks on the Preparation of Argon. V. Vapor Pressure Above -140° C, Critical Temperature and Critical Pressure of Argon," by C. A. Crommelin, Communications from the Physical Laboratory of the University of Leiden, Holland, no. 115, 1910.

26 "Some of the Thermodynamic Properties of Ammonia. I. The Compressibility of and an Equation of State for Gaseous Ammonia. The Vapor Pressure of Liquid Ammonia," by J. A. Beattie and C. K. Lawrence, *Journal of the American Chemical Society*, vol. 52, 1930, p. 6.

27 "The Pressure-Volume-Temperature Values for Ammonia to One Thousand Atmospheres From 30 to 200° ," by F. G. Keyes, *Journal of the American Chemical Society*, vol. 53, 1931, p. 965.

28 "Compressibility of Ammonia at High Pressures and Temperatures," by Ya. S. Kazarnovski, *Acta Physicochimica, USSR*, vol. 12, 1940, p. 513.

29 "The Thermodynamic Properties of Dichlorodifluoromethane, a New Refrigerant, The Equation of State of Superheated Vapor," by R. M. Buffington and W. K. Gilkey, *Industrial and Engineering Chemistry*, vol. 23, 1931, p. 254.

30 "Thermodynamic Properties of Dichlorodifluoromethane (F-12)," Circular No. 12, The American Society of Refrigerating Engineers, 1931.

31 "Thermodynamic Properties of Fluorochloromethanes and Ethanes, P-V-T Relations of Three Fluorochloromethanes and Trifluorotrchloroethanes," by A. F. Benning and R. C. McHarness, *Industrial and Engineering Chemistry*, vol. 32, 1940, p. 698.

32 "On Evaporation and Dissociation, Part II. A Study of the Thermal Properties of Alcohol," by W. Ramsay and S. Young, *Philosophical Transactions of the Royal Society, London, England*, series A, vol. 177, 1886, p. 123.

33 "The Pressure-Volume-Temperature Relation for Gaseous Ethyl Ether I," by J. A. Beattie, *Journal of the American Chemical Society*, vol. 46, 1924, p. 342.

34 "The Pressure-Volume-Temperature Relation for Gaseous Ethyl Ether," by J. A. Beattie, *Journal of the American Chemical Society*, vol. 49, 1927, p. 1123.

35 "On Evaporation and Dissociation—Part III. A Study of the Thermal Properties of Ethyl Oxide," by W. Ramsay and S. Young, *Trans. of the Royal Society, London, England*, series A, vol. 178, 1887, p. 57.

36 "Volumetric Behavior of Hydrogen Sulphide," by H. H. Reamer, B. H. Sage, and W. N. Lacey, *Industrial and Engineering Chemistry*, vol. 42, 1950, p. 140.

37 "On the P-V-T Relations and an Equation of State for Unsaturated Vapor of Methyl Alcohol," by L. M. Lashakov, *Acta Physicochimica*, vol. 11, 1939, p. 107.

38 "On Evaporation and Dissociation. Part V. A Study of the Thermal Properties of Methyl Alcohol," by W. Ramsay and S. Young, *Philosophical Transactions of the Royal Society, London, England*, series A, vol. 178, 1887, p. 313.

39 "Mésures Concernant la Surface de Van der Waals pour les Melanges d'Acide Carbonique et de Chlorure de Methyle," by J. D. Kuenen, Archives, *Neerlandaises des Sciences, Exact et Naturelles*, vol. 26, 1893, p. 368.

40 "Thermodynamic Properties of Methyl Chloride," by H. G. Tanner, A. F. Benning, and W. F. Mathewson, *Industrial and Engineering Chemistry*, vol. 31, 1939, p. 878.

41 "Isotherms and Thermodynamic Functions of Methyl Fluoride at Temperatures Between 0° and 150° C and at Pressures Up to 150 Atmospheres," by A. Michels, A. Visser, R. J. Lunbeck, and G. J. Wolters, *Physica*, vol. 18, 1952, p. 114.

42 "Volumetric Behavior of Propene," by P. S. Farrington and B. H. Sage, *Industrial and Engineering Chemistry*, vol. 41, 1949, p. 1734.

43 "Compressibility and Critical Constants of Propylene Vapor," by H. Marchman, H. W. Prengle, and R. L. Motard, *Industrial and Engineering Chemistry*, vol. 41, 1949, p. 2658.

44 "The Specific Volume of Steam in the Saturated and Superheated Condition Together With Derived Values of the Enthalpy, Entropy, Heat Capacity and Joule Thomson Coefficients," by F. G. Keyes, L. B. Smith, and H. T. Gerry, *Proceedings of the American Academy of Arts and Sciences*, vol. 70, 1936, p. 319.

45 "Die thermischen, Eigenschaften des gesättigten und des überhitzten wasserdampfes zwischen 100° und 180° C," by O. Knoblauch, C. von Linde, and H. Klebe, *Mitteilungen über Forschungsarbeiten auf dem Gebiete des Ingenieurwesens*, vol. 21, 1905, p. 33.

46 International Critical Tables, McGraw-Hill Book Company, Inc., New York, N. Y., 1926.

47 "Thermodynamic Properties of Dichlorodifluoromethane, A New Refrigerant," by F. R. Bichowsky and W. K. Gilkey, *Industrial and Engineering Chemistry*, vol. 23, 1931, p. 366.

48 "Thermodynamic Properties of Fluorochloromethanes and Ethanes," by A. F. Benning and R. C. McHarness, *Industrial and Engineering Chemistry*, vol. 32, 1940, p. 814.

49 "The Critical Constants and Vapor Pressures at High Pressure of Some Gaseous Fluorides of Group IV," by H. S. Booth and C. F. Swinehart, *Journal of the American Chemical Society*, vol. 57, 1935, p. 1337.

50 "Taschenbuch für Chemiker und Physiker," Springer Verlag, Berlin, Germany, 1949.

51 "P-V-T Relations of Propylene," by W. E. Vaughan and N. R. Graves, *Industrial and Engineering Chemistry*, vol. 32, 1940, p. 1252.

52 "Two-Phase Equilibrium in Binary and Ternary Systems, IV. The System Ethane-Propylene," by H. Lu, D. G. Newitt, and M. Ruhemann, *Proceedings of the Royal Society of London, England*, series A, vol. 178, 1941, p. 506.

53 "Thermodynamic Properties of Steam," by J. H. Keenan and F. G. Keyes, John Wiley & Sons, Inc., New York, N. Y., 1936.

54 "Tables of Selected Properties of Ammonia," Circular of the Bureau of Standards, no. 142, U. S. Department of Commerce, 1923.

55 "Thermodynamic Properties of Dichlorodifluoromethane, a New Refrigerant," by W. K. Gilkey, F. W. Gerard, and M. E. Bixler, *Industrial and Engineering Chemistry*, vol. 23, 1931, p. 364.

56 "Thermodynamic Properties of Freon-21, Dichlorodifluoromethane," by A. F. Benning and R. C. McHarness, Bulletin, Kinetic Chemicals, Inc., 1939.

57 "The Thermodynamic Properties of Freon-22, Monochlorodifluoromethane," by D. P. Graham and R. C. McHarness, Bulletin, Kinetic Chemicals, Inc., 1945.

58 "Thermodynamic Properties of Freon-11, Trichlorofluoromethane," by A. F. Benning and R. C. McHarness, Bulletin, Kinetic Chemicals, Inc., 1938.

59 "Thermodynamic Properties of Hydrogen Sulphide," by J. R. West, *Chemical Engineering Progress*, vol. 44, 1948, p. 287.

60 "Thermodynamic Properties of Methyl Alcohol," by J. F. Taylor, Report 9-2, Aeronautical Engineering Department, University of Illinois, 1946.

61 B. T. Crawford and D. E. Mann, Microwave Spectroscopy Section of Annual Review, *Physical Chemistry*, vol. 1, 1950, p. 151.

62 "A Table of Dipole Moments," by G. C. Hampson and R. J. B. Marsden, *Transactions of the Faraday Society*, Appendix, vol. 30, 1934.

63 "Dielectric Constant and Molecular Structure," by C. P. Smyth, Chemical Catalogue Company, New York, N. Y., 1931.

64 "Molecular Microwave Spectra Tables," by P. Kisliuk and C. H. Townes, *Journal of Research*, U. S. National Bureau of Standards, vol. 44, 1950, p. 611.

65 "An Extension of the Law of Corresponding States," by W. E. Ibele, PhD thesis, University of Minnesota, August, 1953.

66 "The Nature of the Chemical Bond," by L. Pauling, Cornell University Press, Ithaca, N. Y., 1940.

67 "Thermische und kalorische Messungen an sehr reinem

Schwefeldioxyd," by L. Riedel, *Zeitschrift für die Gesamte Kälte-Industrie*, vol. 46, 1939, p. 22.

68 "An Empirical Equation for Thermodynamic Properties of Light Hydrocarbons and Their Mixtures," by M. Benedict, G. B. Webb, and L. C. Rubin, *The Journal of Chemical Physics*, vol. 8, 1940, p. 334.

Discussion

K. S. PITZER.⁴ This paper is very interesting to the writer because we are just completing a somewhat similar study of compressibility deviations for nonpolar molecules of irregular shape. The present study on polar gases should supplement our work. The selection of the reduced dipole moment parameter appears to be sound theoretically and the resulting correlations are very successful for molecules of compact and nearly spherical shape, but with increasing dipole moments. The deviations shown in Figs. 2 and 3 for elongated molecules of low dipole moment are to

⁴ Dean, College of Chemistry, University of California, Berkeley, Calif.

be understood in terms of the theory which we are proposing. The only other molecules showing significant deviation are methyl alcohol and ethyl alcohol. In these cases the dipole moment is concentrated in a portion of the molecule and it is not surprising that the behavior is more complex than for the other examples.

AUTHORS' CLOSURE

The authors are pleased to acknowledge the discussion of Dean Pitzer. One additional comment regarding experimental P - V - T data would appear to be in order. Early in the course of the study, it soon became apparent that despite the abundance of gases involved in industrial processes, accurate information in so far as number of gases involved and/or the pressure and temperature range covered was quite limited. While the results presented here represent an effort to improve the status of gas-property knowledge in a very real sense, it also indicates a continuing need for experimental data covering wider ranges of pressure and temperature for pure gases and mixtures.

Total Normal Emissivity Measurements on Aircraft Materials Between 100 and 800 F

By N. W. SNYDER,¹ J. T. GIER,² AND R. V. DUNKLE,³ BERKELEY, CALIF.

The results of techniques developed and used to determine the mean effective emissivity of different surfaces over a range of temperatures from 100 to 800 F are presented. The measurement of radiant energy from the surfaces is accomplished by means of a 160-junction silver-constantan directional thermopile radiometer. Serious errors are noted and accounted for when improper use is made of polished shields in the experimental arrangements. This represents a refinement over previous techniques used; further refinements are anticipated with the ultimate aim of checking against integrated spectral data. However, the data presented are useful aids as a phase of the technique of evaluation of equilibrium temperatures of different surfaces exposed to solar irradiation.

NOMENCLATURE

The following nomenclature is used in the paper:

- A_1, A_2 = areas of two surfaces radiating directly to each other, sq ft
 A_0 = area of sample surface seen by thermopile receiver, sq ft
 A_R = area of thermopile receiver surface, sq ft
 A_H = area of heater case surrounding sample, sq ft
 $B = \frac{F_{H,R}}{F_{0,R}} \frac{e_H(1 - e_X)}{(1 - Z)}$, a factor, dimensionless
 $C = \frac{T_H^4 - T_R^4}{T_0^4 - T_R^4}$ = ratio of temperature differences, dimensionless
 C_0 = heat conduction to supports for a sample exposed to atmospheric radiation per unit area of sample, Btu/hr sq ft
 e_{me} = total emissivity of a surface at a given temperature of surface; named mean effective emissivity
 e_λ = monochromatic emissivity of a surface at temperature T
 e_{lw} = emissivity of surface with respect to plate surface at low temperatures (long-wave radiation)
 $e_0 = e_{me(0)}$ = mean effective emissivity of sample test surface
 $e_g = e_{me(g)}$ = incorrect mean effective emissivity of sample test surface because of test procedure
 e_X = emissivity of shield
 e_H = emissivity of heater case
 e_s = emissivity of surroundings

- E = emissive power of a surface at temperature T , Btu/hr sq ft
 E_λ = monochromatic power of a surface at temperature T , Btu/hr sq ft/(micron)
 $E_{\lambda i}$ = monochromatic emissive power of an ideal radiator at temperature T , Btu/hr sq ft (micron)
 $F_{1,2}$ = shape factor of surface 1 with respect to 2 where surface 1 is being irradiated by 2, dimensionless
 $F_{2,1}$ = vice versa of $F_{1,2}$
 $F_{0,s}$ = shape factor of sample surface with respect to surrounding surfaces (conversely for $F_{s,0}$)
 $F_{R,0}$ = shape factor of thermopile receiver surface with respect to sample surface (conversely for $F_{0,R}$)
 $F_{R,H}$ = ratio of all incident energy upon receiver which is emitted by heater case to total energy emitted by heater case
 $F_{R,X,H}$ = ratio of radiant energy incident upon receiver from heater case after one reflection from shield and sample surface to total energy from heater case
 $F_{H,X,H}$ = ratio of radiant energy incident upon heater case from heater case after one reflection from shield to total energy from heater case
 G = unit incident radiant energy upon a surface, Btu/hr sq ft
 G' = total incident radiant energy upon a surface, Btu/hr
 G_{sw} = short-wave irradiation from sun plus sky, Btu/hr sq ft
 G_{lw} = long-wave irradiation from sun plus sky, Btu/hr sq ft
 G_1 = irradiation of thermopile receiver by energy emitted from sample surface 0, Btu/hr sq ft
 G_s = irradiation of thermopile receiver by energy emitted by surroundings and reflected directly from sample surface, Btu/hr sq ft
 G_H = irradiation of thermopile receiver by energy emitted by heater case and reflected by both shield and sample surface, Btu/hr sq ft
 G_{H0} = irradiation of thermopile receiver by energy from heater case after one reflection from shield and sample, Btu/hr sq ft
 G_{H1} = irradiation of thermopile receiver by energy emitted by heater case after two reflections from shield and one from heater and sample, Btu/hr sq ft
 G_{Hn} = energy which is multiply reflected between heater case and shield and then from sample onto thermopile receiver = $G_{H0}, G_{H1}, G_{H2}, \dots, \sum_{n=0}^{\infty} G_{Hn} = G_H$, Btu/hr sq ft; n = number of reflections
 h_c = heat-transfer coefficient for convection from surface to air, Btu/hr sq ft deg F
 K = radiometer sensitivity constant, Btu/hr sq ft (Mv)
 L = distance from thermopile receiver to opening in shield, ft
 Mv = thermoelectric output of thermopile, millivolts
 $Q = e_0 F_{0,R} \sigma (T_0^4 - T_R^4)$, Btu/hr sq ft
 $Q_{net} = A_R e_{me} F_{R,0} \sigma (T_0^4 - T_R^4)$, net energy interchange be-

¹ Associate Professor, Mechanical Engineering Department, University of California.

² Associate Professor of Electrical Engineering, University of California.

³ Associate Professor of Mechanical Engineering, University of California.

Contributed by the Heat Transfer Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 29, 1954. Paper No. 54-A-189.

tween thermopile receiver and sample surface,
Btu/hr
 r = radius of opening in shield, ft
 t_0 = temperature of sample surface, deg F
 t_{air} = temperature of ambient air, deg F
 $T_0 = t_0 + 460$, deg R
 T_R = absolute temperature of thermopile receiver, deg R
 T_S = absolute temperature of surroundings, deg R
 T_H = absolute temperature of heater case, deg R
 $Z = F_{H.X.H}(1 - e_X)(1 - e_H)$, dimensionless
 α_{sw} = absorptivity of surface with respect to short-wave solar plus sky irradiation
 α_{lw} = absorptivity of surface with respect to long-wave solar plus sky irradiation
 $\gamma = e_H(1 - e_X)(1 - e_0) \left(\frac{1}{1 - Z} \right) F_{H.R} (T_H^4 - T_R^4)$,
Btu/hr sq ft
 λ = wave length, μ , microns
 μ = microns

INTRODUCTION

This paper represents one phase of the experimental work in connection with the development of techniques for the prediction of equilibrium temperatures for surfaces exposed to solar irradiation (1).⁴ It may be looked upon as the first step in the evaluation of the spectral properties of materials over broad "long wave-length" regions.

Surfaces that are selective absorbers or reflectors will manifest this selectivity by attaining different equilibrium temperatures when exposed to irradiation from sources such as the sun which differs appreciably from a uniform energy spectrum. For instance, when irradiated by the sun, surfaces which have a low absorptivity for "short wave-length" irradiation and a high long wave-length absorptivity (or emissivity) will be cooler (attain a lower equilibrium temperature) than surfaces of high short wave-length absorptivity and low long wave-length emissivity.

The ratio of the short wave-length emissivity to the long wave-length emissivity of any given materials can be used as indexes for predicting relative equilibrium temperatures for the materials when exposed to any given solar and atmospheric irradiation.

It is presumed that the surfaces are being irradiated by both solar and sky energy. For convenience, it is desirable to differentiate between short wave-length energy as $\lambda < 1\mu$ and long wave-length energy as $\lambda > 1\mu$. Most of the solar energy lies in the short wave-length region and most of the energy from a 200 to 300 F surface lies in the long wave-length region. The heat balance for a plate exposed to solar irradiation may be expressed as follows

$$\alpha_{sw}G_{sw} + \alpha_{lw}G_{lw} = e_{lw}\sigma T_0^4 + h_c(t_0 - t_{air}) + C_0 \dots [1]$$

where

G_{sw} = short-wave irradiation from sun plus sky, Btu/hr sq ft
 G_{lw} = long-wave irradiation from sun plus sky, Btu/hr sq ft
 α_{sw} = absorptivity of plate surface with respect to short-wave solar plus sky irradiation
 α_{lw} = absorptivity of plate surface with respect to long-wave solar plus sky irradiation
 e_{lw} = emissivity of plate surface with respect to radiation from plate surface at low temperatures
 σ = Stefan-Boltzmann constant (1.714×10^{-9} Btu/hr sq ft deg R⁴) (reference 6)
 t_0 = surface temperature of plate, deg F

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

h_c = heat-transfer coefficient for convection from plate, Btu/hr sq ft deg F
 $T_0 = t_0 + 460$, absolute temperature, deg R
 t_{air} = ambient temperature of air, deg F
 C_0 = heat conduction to supports for plate, Btu/hr sq ft

The value of α_{sw} may be determined in Equation [1] provided that G_{sw} , G_{lw} , α_{lw} , e_{lw} , h_c , and C_0 are known. C_0 may be made negligible; h_c may be determined from existing heat-transfer data; α_{lw} may be considered equal to e_{lw} and determined experimentally; G_{sw} and G_{lw} may be determined experimentally. Techniques for evaluating e_{sw} have been discussed in reference (1).

There is a certain amount of data for e_{lw} available (1) but the techniques used in the evaluation have not been described completely in the original references and certain errors in the methods of mensuration appear to be evident, based upon data and analysis in this paper. Omission of much data is shown in the tables of reference (1). The following is a result of the investigation on the mensuration techniques and determinations of the different values of e_{lw} for a few surfaces. Although accurate data for e_{lw} were desired for temperatures at approximately 100 F or even lower, they were found difficult to obtain with the techniques used because of the small amounts of net energy available for measurement. It was concluded that data for slightly higher temperatures could be extrapolated to the lower temperatures, provided the variation of emissivities was somewhat linear with temperature. The symbol e_{ms} (mean effective emissivity at surface temperature T) was used instead of e_{lw} and will be explained in greater detail in the next section. These mean effective values are based on certain idealizations and further refinements are anticipated. It is believed that more accurate results may be obtained by integrating a spectroradiometric curve obtained by use of an infrared Perkin-Elmer spectrometer now available in the laboratory. Prisms can be used for measurements to 25μ .

The technique of measurement employed for determining e_{ms} was first used in this laboratory in 1941 (2). That project was carried out in co-operation with the National Advisory Committee for Aeronautics. Some errors inherent in the technique that was used have been detected and accounted for. New equipment has been designed to minimize these errors in the measurements.

Tested surfaces were aluminum, stainless steel, magnesium, and various external paints and lacquers.

DEFINITION OF MEAN EFFECTIVE EMISSIVITIES

The radiant energy emitted from an ideal (Planckian) radiator may be expressed by the Stefan-Boltzmann equations as

$$E_i = \sigma T^4 \dots [2]$$

where

E_i = emissive power from an ideal radiator at temperature T , Btu/hr sq ft
 σ = Stefan-Boltzmann constant, 1714×10^{-12} Btu/hr sq ft deg R⁴
 T = absolute temperature of surface, deg R

By designating $E_{\lambda i}$ as the radiant energy emitted per unit wave-length at the wavelength λ for an ideal radiator, the emissive power may then be represented as

$$E_i = \int_0^\infty E_{\lambda i} d\lambda \dots [3]$$

For a nonideal radiator, the emissive power becomes

$$E = \int_0^\infty e_\lambda E_{\lambda i} d\lambda \dots [4]$$

where e_λ is the ratio of ideal radiation to actual radiation for the

same surface temperature at a given wave length and is designated as the monochromatic emissivity.

The value of the monochromatic emissivity, ϵ_λ , should really be considered as a function of the direction of radiation with respect to the normal of the emitting surface. If E is then construed to be the emitted energy from the surface over a hemisphere, it therefore must be based upon a mean value of ϵ_λ with angle or upon the normal value. The measurements made for determining E were taken over a small region about the normal. An idealization is made that the value of ϵ_λ is constant and the surface perfectly diffused over this angle. By dividing the emitted energy from the surface E , with respect to measurements made at the normal by the ideal radiation, a mean effective value of the emissivity with respect to wave length at the normal is obtained.

Now let

$$\epsilon_{me} = \frac{E}{E_i} = \frac{\int_0^\infty \epsilon_\lambda E_\lambda d\lambda}{\int_0^\infty E_\lambda d\lambda} \quad [5]$$

Equation [5] may be considered to define the mean effective emissivity ϵ_{me} , with respect to radiation from a surface at any angle, the mean value with respect to angle or the value at the normal.

Equation [4] becomes

$$E = \epsilon_{me} \int_0^\infty E_\lambda d\lambda \quad [6]$$

or

$$E = \epsilon_{me} E_i \quad [7]$$

also, by substituting Equation [2] in Equation [7]

$$E = \epsilon_{me} \sigma T^4 \quad [8]$$

The definition of ϵ_{me} also can be stated by means of Equations [8] and [4] as

$$\epsilon_{me} = \frac{E}{\sigma T^4} = \frac{\int_0^\infty \epsilon_\lambda E_\lambda d\lambda}{\sigma T^4} \quad [9]$$

The value ϵ_{me} is then expressed as the ratio of radiant energy emitted by a surface at the temperature T to the radiation from an ideal radiator at the same temperature. The heat loss from a surface by radiation may be determined with Equation [8] provided ϵ_{me} is known. Several values of ϵ_{me} have been compiled for different surfaces in reference (1).

It was the intention of the experiments to measure values of ϵ_{me} for different surfaces and investigate the errors in measurement. Equation [9] yielded a means of measuring ϵ_{me} by merely determining E and dividing by σT^4 . The value of E as obtained by means of a thermopile radiometer contains certain errors which were discussed previously (2) and are also discussed in this paper.

EQUIPMENT

Samples. The samples tested were 6 in. square and included the types of surfaces listed in Table 1.

Heater. The heater consisted of a heavy copper plate, with electric heating coils on one side, to which the samples were attached. The copper plate served to distribute the heat uniformly to the sample. The electric coils were insulated from the ambient air by means of fire clay and a transite case, as indicated in Fig. 1. Fig. 2 shows the iron-constantan thermocouple mounting. Figs. 3 and 4 show the relative position of the heater in the test arrangement.

Radiometer. The radiant energy was detected by means of a thermopile which has been described in reference (3). Briefly,

TABLE 1 TYPES OF SURFACES OF SAMPLES

Fig. no.	Description	Identification
Painted surfaces		
9.....	W. P. Fuller Co. External air-drying enamel, D-70-6324	1
10.....	Aluminized lacquer 1234 over 75-ST alclad	19
11.....	Clear lacquer 1234 over 75-ST alclad	14
Steel surfaces		
12.....	18-8 stainless steel, sandblasted, weathered	15.1
13.....	18-8 stainless, oxidized at 1500 F, weathered	5.2
14.....	18-8 stainless, chromic, and sulphuric acid treated	5.4
15.....	18-8 stainless, unpolished	15
16.....	18-8 stainless, aerobright polish	16
Aluminum surfaces		
17.....	53-SO aluminum alloy, weathered	0.6
18.....	24-ST aluminum alloy, weathered	0.8
19.....	75-ST alclad, unpolished	10
20.....	75-ST alclad, aerobright polish	11
21.....	75-ST alclad, anodized	18
Other metals		
22.....	J. H. magnesium alloy, weathered	7.1
23.....	MH-42 magnesium alloy	7.3

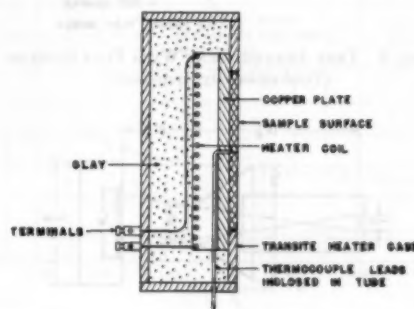


FIG. 1 CROSS SECTION OF HEATER AND ATTACHED SAMPLE SURFACE

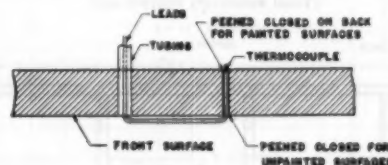


FIG. 2 IRON-CONSTANTAN THERMOCOUPLE MOUNTING ON SAMPLE SURFACE

the thermopile consisted of approximately 104 turns per inch of No. 40 AWG constantan wire wound over glass insulators which were fastened to a copper spool, Fig. 5. Silver was deposited on a portion of the coil in such a manner as to leave two lines of silver-constantan junctions.

Aluminum-foil strips, lamp-blackened on the exposed side, were fastened to the two lines of junctions with dilute shellac. The cold-junction receiver strip was shielded from radiation that entered the front end of the radiometer by a plate that was at ambient temperature. This shield was lamp-blackened on the side facing the cold junctions and was polished on the side facing the opening of the radiometer. The hot-junction receiver strip was exposed to radiant energy from the sample through a narrow window.

For the purpose of maintaining stability and a directional characteristic, the thermopile was mounted in a metal case. Fig. 5 shows the general assembly. The sensitivity of the radiometer, utilized for the tests described in this paper, was $K = 7.84$ Btu/hr sq ft (Mv). The output of the thermopile was measured by means of a Leeds and Northrup No. 8662 potentiometer. Further refinement involved using a type K potentiometer. The radiometer was calibrated against a standard tungsten-filament lamp which previously had been calibrated by the Bureau of Standards.

Shields. A shield was placed between the radiometer and sample so as to prevent direct radiation from the heater case from striking the thermopile. The first shield used consisted of five polished 24-ST aluminum plates with a $3\frac{3}{8}$ -in. circular opening placed $6\frac{1}{2}$ in. from the sample, Fig. 3. It was found that radiant energy from the heater case reflected from the shield onto the sample and then into the radiometer, causing considerable error when the emissivity of the sample was less than approximately

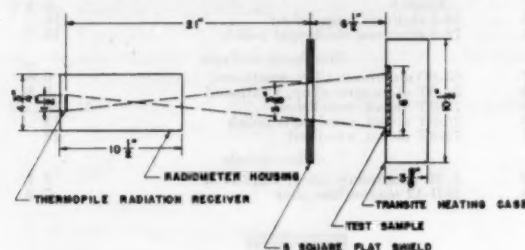


FIG. 3 TEST ARRANGEMENT WITH FLAT SHIELDS
(Total emissivity equipment.)

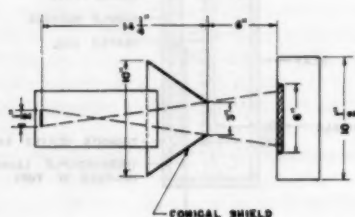


FIG. 4 TEST ARRANGEMENT WITH CONICAL SHIELD
(Total emissivity equipment.)

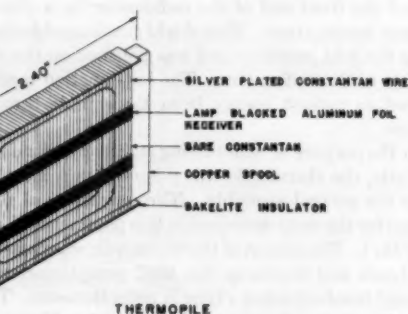
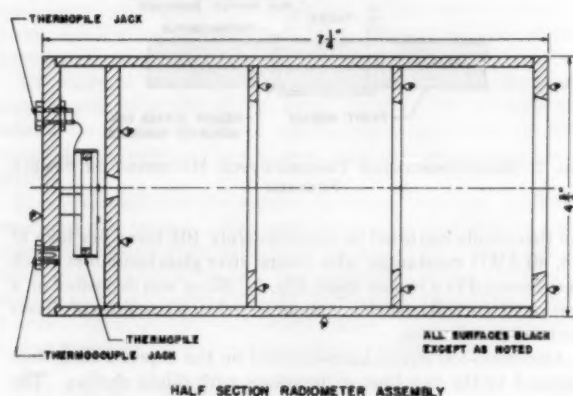


FIG. 5 THERMOPILE RADIOMETER

0.4. One interesting test was made by lamp-blackening the shield in order to reduce these reflections. A redesign of the shield led to an arrangement indicated in Fig. 4. The shield was shaped into a truncated cone to eliminate radiant energy emitted by the heater case from being reflected from the shield and sample into the radiometer.

The error of using a flat shield is particularly emphasized in this paper.

PROCEDURE

The samples were maintained at a constant temperature by means of the heater, and the surface temperatures were measured with an iron-constantan thermocouple. It was found that the temperatures were uniform over the entire surface so that one thermocouple was used and was mounted in a fashion indicated in Fig. 2.

A solid shield was placed between the radiometer and heated sample until a reading was to be taken. A number of readings were taken at approximately the same sample temperature. This process was repeated over a range of temperatures for each of the samples.

DATA AND COMPUTATIONS

Computations With Flat Shield. From Equation [9] the value of E was determined by measuring the emitted energy from a surface with a radiometer. Reference (5) presents an analysis of the meaning of the results obtained with the radiometer of which Equation [10] is a result

$$\frac{q_{\text{net}}}{A_R} = e_{\text{ms}} F_{R,0} \sigma (T_0^4 - T_R^4) \dots \dots \dots [10]$$

but q_{net}/A_R is equal to the energy detected by the radiometer as $K(\text{Mv})$. Therefore (see references 2 and 4)

$$e_{\text{ms}} = \frac{K(\text{Mv})}{F_{R,0} \sigma (T_0^4 - T_R^4)} \dots \dots \dots [11]$$

From Fig. 3, $F_{R,0}$ (the shape factor) was determined by the expression

$$F_{R,0} = \frac{r^2}{r^2 + L^2} \dots \dots \dots [12]$$

$$F_{R,0} = \frac{(3.38/2)^2}{(3.38/2)^2 + (21)^2} = 0.00642$$

Equation [11] becomes

$$e_{\text{ms}} = \frac{7.84 (\text{Mv})}{(0.00642)(0.1714) \left[\left(\frac{T_0}{100} \right)^4 - \left(\frac{T_R}{100} \right)^4 \right]} \dots \dots [13]$$

$$e_{\text{ms}} = \frac{7120 (\text{Mv})}{\left[\left(\frac{T_0}{100} \right)^4 - \left(\frac{T_R}{100} \right)^4 \right]}$$

Data obtained with the flat shield and utilized in Equation [13] are indicated as such in Figs. 9, 14, and 23. The error of using a flat shield here was considered negligible because of the relatively high emissivity of the surfaces tested.

Computations With Conical Shield. The arrangement of the shield and radiometer is shown in Fig. 4. Then

$$F_{R,0} = \frac{r^2}{r^2 + L^2}$$

$$F_{R,0} = \frac{(3/2)^2}{(3/2)^2 + (14\frac{1}{4})^2} = 0.011$$

and Equation [11] becomes

$$\epsilon_{me} = \frac{7.84 \text{ (Mv)}}{(0.011)(0.1714) \left[\left(\frac{T_0}{100} \right)^4 - \left(\frac{T_R}{100} \right)^4 \right]}$$

or

$$\epsilon_{me} = \frac{4160 \text{ (Mv)}}{\left[\left(\frac{T_0}{100} \right)^4 - \left(\frac{T_R}{100} \right)^4 \right]} \quad [14]$$

The data taken with the conical-shield arrangement are represented in Figs. 10 through 13 and 15 through 22.

DISCUSSION OF ERRORS

Calibration of Radiometer. The radiometer was calibrated by means of a high-intensity tungsten-filament lamp which emitted energy predominantly in the near infrared. Because the bulb of the lamp was made of glass much of the far infrared was absorbed. However, a small net amount of infrared energy was emitted from the glass bulb because of its temperature being somewhat above the ambient air. By calibrating a thermopile with this type of source and then using the thermopile to measure accurately radiant energy possessing a considerable percentage of far infrared would require that the absorptivity of the receiver surface have a flat characteristic with respect to wave length throughout the visible and near infrared.

The surface treatment of the thermopiles consisted of a lamp-black coating. It is known that lampblack is not only a good absorber but also possesses a constant absorptivity with wave length in the near infrared. However, the absorptivity of lampblack decreases above 15μ . An ideal radiator surface, at, say, 200 F, yields a peak energy output at 7.9μ but still emits approximately 25 per cent of the total radiant energy above 15μ . This shows that it is important to calibrate the thermopile with radiant energy from a source which emits in the same relative magnitudes with respect to wave length as the test surface. Since this is not entirely possible, investigations are now in progress to furnish methods for determining the emissivities by means of an infrared spectrometer.

A study also will be made to determine quantitatively the accuracy and limitations of the radiometric method for obtaining mean effective emissivities (as described in this paper). This will be done by making a spectroradiometric investigation of the absorptivity of the thermopile receiver and the emissive power of the calibration source.

Measurement of Radiation From Sample Surface. The first measurements made were obtained with the flat-shield arrangement shown in Fig. 3. The techniques used were the same as those indicated in references (2) and (4). In computing the emissivities by means of Equation [13], no correction was made for the energy emitted from the heater surface which surrounded the sample surface and which was reflected from the shield to the sample and then onto the thermopile. This error was determined by independent methods: (a) the side of the flat shield facing the sample was lampblacked in order to reduce the reflected energy to zero, and (b) the shield was shaped conically, Fig. 4, so as to reflect the unwanted energy other than toward the sample surface. These errors were found to be large for sample surfaces with high reflectivities. Fig. 6 shows the results of data taken by the foregoing methods with a polished 75-ST Alclad sample surface. The emissivities determined with the conical and lampblacked flat shield were practically the same while those with the flat polished shield were from 4 to 5 times higher.

Fig. 7 indicates a comparison of data obtained for a weathered

24-ST aluminum-alloy sample surface using the conical shield and the polished flat shield. This comparison shows that lower or negligible errors may be encountered with a more diffuse surface of higher emissivity. In Fig. 8 the data for a weathered 53-SO aluminum-alloy surface shows an even higher emissivity with no difference between the data taken with the polished flat shield and the conical shield.

The Appendix presents an analysis which was made for estimating the foregoing error with the polished flat shield. The results indicated that, Equation [48]

$$e_0/e_g = 1 - \left[\frac{BC}{\left(\frac{e_g}{1-e_0} \right) + BC} \right] \quad [48]$$

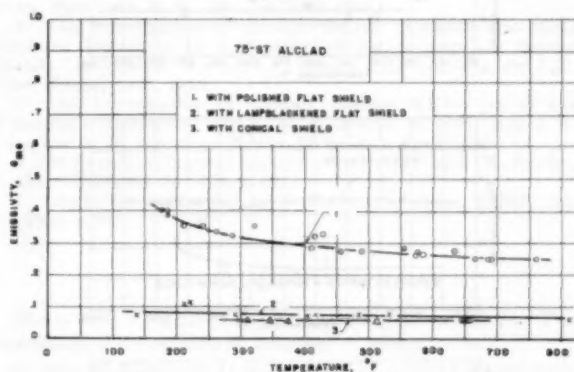


FIG. 6

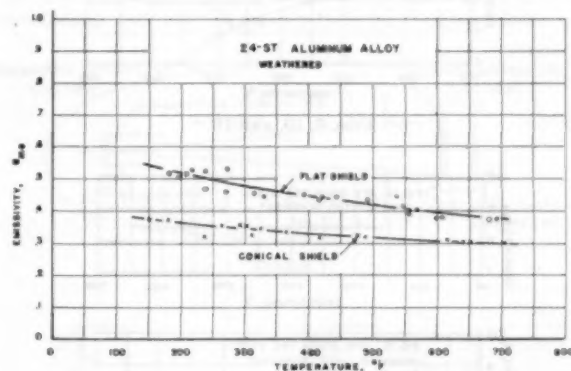


FIG. 7

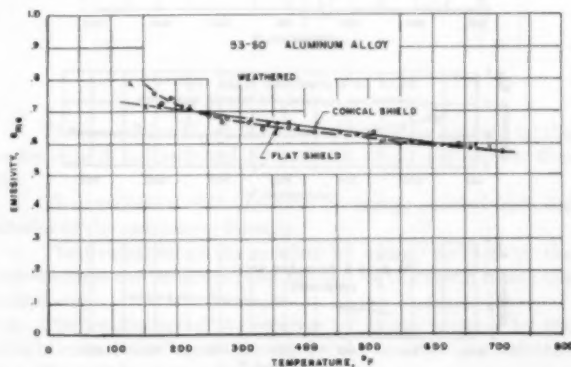


FIG. 8

where

e_0 = correct e_{ms} of sample surface

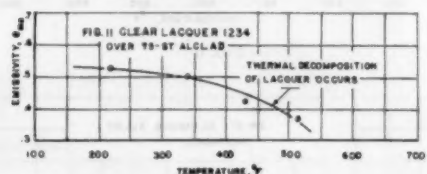
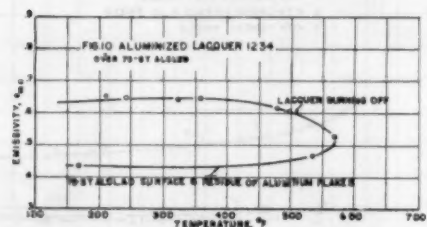
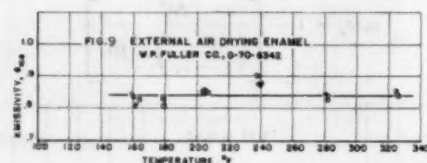
e_p = incorrect e_{ms} of sample surface because of additional energy added by heater case through multiple reflections

$$BC = \frac{F_H R}{F_{e,R}} \frac{e_H(1 - e_x)}{1 - Z} \frac{(T_H^4 - T_R^4)}{(T_0^4 - T_R^4)}, \text{ an interreflection factor}$$

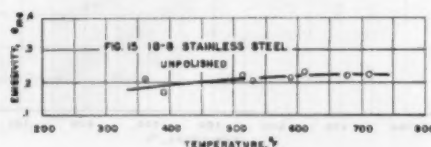
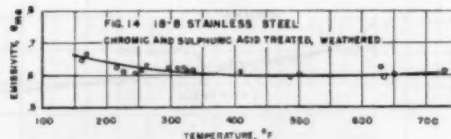
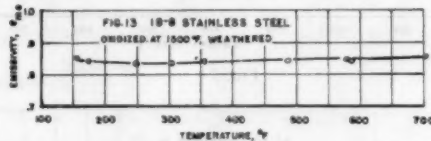
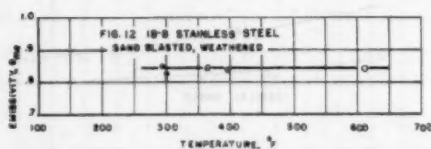
Equation [48] shows that e_p , the incorrect emissivity, ap-

proaches e_0 , the correct emissivity, as $e_0 \rightarrow 1$ with other variables remaining constant (those included in BC). When e_0 is small the inaccuracy of e_p becomes large because

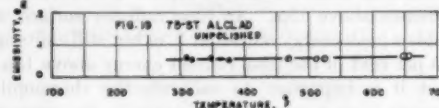
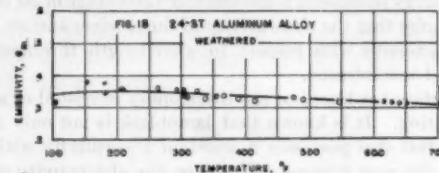
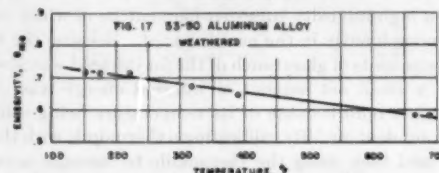
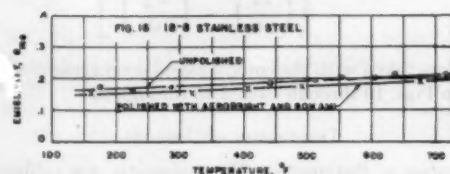
$$\frac{BC}{e_0 BC} \rightarrow 1$$



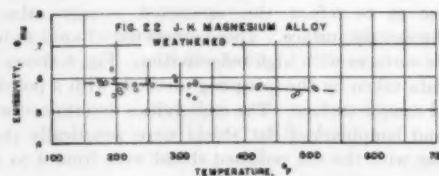
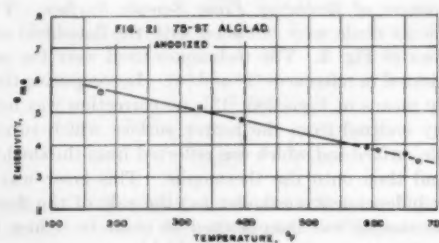
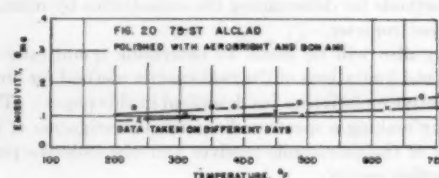
FIGS. 9, 10, AND 11



FIGS. 12, 13, 14, AND 15



FIGS. 16, 17, 18, AND 19



FIGS. 20, 21, AND 22

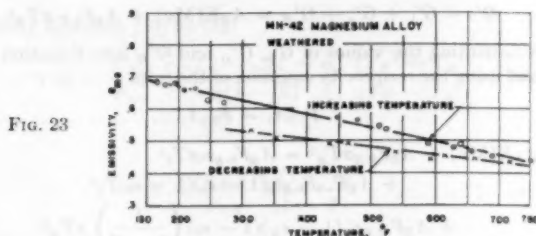


Fig. 23

With e_0 small but finite, the error can still be reduced to negligible value by making the shield "black" ($e_x \cong 1$) which, in effect, reduces the term BC toward zero. This was noted previously in connection with data taken for Fig. 6.

After making the foregoing analysis and experiments it was decided to use only a conical shield for all the future tests. The results shown in Figs. 10 through 13, and 15 through 22, were those taken with the conical shield. Some of the results taken with the flat shield were used because the emissivities were high enough to yield few of the errors mentioned.

Slope of Curves. A consideration of the theoretical results for the reflection of radiation from metals (7) shows that the slope of e_{me} versus temperature should be positive. Opaque oxides or dielectrics exhibit a negative slope. Experimental data for polished metals are generally coincident with theoretical predictions (in the infrared region) because the oxide films are thin and relatively transparent. Where the metal surface is weathered or coated, its characteristic is that of the heavy oxide or coating. Thus negative slopes on the e_{me} versus t curves are to be expected when the surface is other than a polished surface.

Note that aluminum, which is easily weathered, shows a positive slope when clean and a negative slope when weathered. Stainless steel, however, with a stable tough oxide, possesses a positive e_{me} versus t slope after relatively harsh treatment. When the treatment amounts to exposure with chromic-sulphuric acid, the stainless steel finally develops a negative e_{me} versus t slope.

It was interesting to note the effect of using an aluminized lacquer as against a clear lacquer in Figs. 10 and 11. The results show that the organic carrier with its strongly absorbing characteristics in the infrared region has a predominant effect.

FURTHER RESEARCH

1 An evaluation of errors involved in using the thermopile for measuring total radiation when calibrated with a source whose spectral emissive power is not the same as the sample surface. This evaluation consists of:

- (a) Determination of the transmissivities of various "blackened" receiver surfaces.
- (b) A determination of the emissive power as a function of wave length for standard sources such as the tungsten-filament lamp.

The thermopile receivers also will be calibrated with the ideal radiators now being developed (1).

2 A determination of the emissivities as a function of angle of the plate with respect to the line of measurement.

3 A complete program of determining the reflectivities and absorptivities of various surfaces by means of a Perkin-Elmer spectrometer to 25μ , now under way.

ACKNOWLEDGMENTS

Acknowledgments are extended to Mr. David Turner, a former member of the laboratory staff, for collecting most of the data and for performing the initial analysis of errors which is presented in revised form in the Appendix; and to Mr. R. Bromberg of the Engineering Department of the University of California at Los An-

geles, for his suggestions and helpful comments regarding the experimental techniques.

The materials under study were samples supplied by the Douglas Aircraft Company, Santa Monica, Calif.

The Office of Naval Research supported the project.

BIBLIOGRAPHY

- 1 Quarterly Progress Report No. 1, Report No. 2, N7-onr-295-Task I, August 12, 1947.
- 2 "An Investigation of Aircraft Heaters XV—The Emissivity of Several Materials," by L. M. K. Boelter, R. Bromberg, and J. T. Gier, NACA ARR, January, 1944.
- 3 "The Silver Constantan Plated Thermopile Temperature, Its Measurement and Control in Industry," by J. T. Gier and L. M. K. Boelter, American Institute of Physics, Reinhold Publishing Company, New York, N. Y., 1941, pp. 1284-1292.
- 4 "An Investigation of Aircraft Heaters XXXII—The Mean Effective Emissivity as a Function of Temperature for Several Metals," by L. M. K. Boelter, J. C. Brannan, R. Bromberg, and J. T. Gier, NACA, ARR, 1945.
- 5 "An Investigation of Aircraft Heaters XXV—Use of the Thermopile Radiometer," by L. M. K. Boelter, R. Bromberg, J. T. Gier, and E. R. Dempster, NACA ARR 5A13, April, 1945.
- 6 "A Review of Thermal-Radiation Constants," by N. W. Snyder, Trans. ASME, vol. 76, 1954, p. 537.
- 7 "Radiation in Metals," by N. W. Snyder, Trans. ASME, vol. 76, 1954, p. 541.

Appendix

The system is described by the diagram given in Fig. 24. The irradiation of the receiver is composed of several parts:

- 1 The irradiation G_1 of the receiver by emitted energy from the sample surface 0.
- 2 The irradiation G_2 of the receiver by energy emitted by the surroundings and reflected directly from the sample surface.

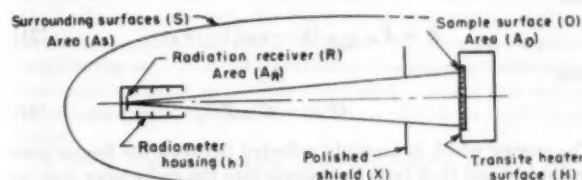


Fig. 24

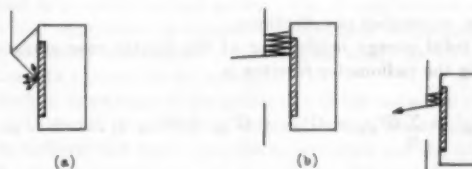


Fig. 25

Fig. 26

3 The irradiation G_B of the receiver by energy emitted by the heater case H and reflected by both the shield and sample, Fig. 25.

4 The irradiation of the receiver by energy emitted from the interior of the radiometer housing.

5 The irradiation of the receiver by energy emitted by the surroundings and reflected from the side of the shield facing the radiometer.

6 The irradiation of the receiver by energy emitted by the sample surface not "seen" directly by the receiver and reflected from the shield and sample surface, Fig. 26.

Items 5 and 6 will be considered negligible and all radiation

will be considered diffuse. This latter assumption is somewhat erroneous but will suffice in that the following analysis is to be used as a qualitative indication of the errors instead of a precise evaluation. Let the total irradiation of the receiver be

$$G'_1 = G_A R \dots [15]$$

and

$$G'_s = G_A R \dots [16]$$

From the Stefan-Boltzmann equation

$$G'_1 = A_0 F_{R,0} \dots [17]$$

$$G'_s = A_s F_{0,s} \dots [18]$$

The value $e_{\text{me}(0)}$ for the sample surface 0 is designated here as e_0 for convenience.

Item 4 has been discussed adequately in a previous report (2), which demonstrated that for the conditions employed, it was negligible (the radiometer housing and receiver were substantially at room temperature). An evaluation of G_H is given by the following method which will not be exact because of the complexity of the system:

Let $G'_{H0} = G_{H0} A_R$ = irradiation of receiver by energy from heater case after one reflection from shield and sample surface. Then

$$G'_{H0} = A_H F_{H,X,H} e_H (1 - e_X) (1 - e_0) \sigma T_H^4 \dots [19]$$

The energy radiated by H which is reflected by X to H then to X , then to 0, and then to R can be approximately expressed as Fig. 25(b)

$$G'_{H1} = G'_{H0} [F_{H,X,H} (1 - e_X) (1 - e_H)] \dots [20]$$

Let

$$Z = F_{H,X,H} (1 - e_X) (1 - e_H) \dots [21]$$

then

$$G'_{H1} = G'_{H0} Z \dots [22]$$

The energy which is multiply reflected between the heater case and shield and then from the sample into the radiometer may be expressed as

$$G'_{Hn} = G'_{H1} Z \dots [23]$$

where n = number of reflections.

The total energy originating at the heater case surface and reaching the radiometer receiver is

$$G'_H = \sum_0^n G'_{Hn} = G'_{H0} + G'_{H1} + G'_{H2} + \dots + G'_{Hn} \dots [24]$$

or

$$G'_H = G'_{H0} (1 + Z + Z^2 + \dots + Z^n)$$

and for $n \rightarrow \infty$

$$G'_H = G'_{H0} \left(\frac{1}{1 - Z} \right) \dots [25]$$

The shape factors $F_{R,X,H}$ and $F_{H,X,H}$ depend upon reflection from the shield which is a specular reflection. Therefore an evaluation of $F_{R,X,H}$ and $F_{H,X,H}$ would be somewhat difficult and may be construed to include the effect of specular reflections.

The total irradiation of the receiver includes the values determined in Equations [17], [18], and [25] and are equated to the energy detected by the radiometer

$$G'_0 = G'_1 + G'_s + G'_H = A_R K(Mv) + A_R F_{0,R} \sigma T_R^4 \dots [26]$$

Substituting the values of G'_1 , G'_s , and G'_H into Equation [28] and using the reciprocity equation of the form

$$F_{1,2} A_2 = F_{2,1} A_1 \dots [27]$$

$$\begin{aligned} A_R K(Mv) + A_R F_{0,R} \sigma T_R^4 &= A_R F_{0,R} e_0 \sigma T_s^4 \\ &+ A_R F_{1,0} F_{0,R} e_s (1 - e_0) (1 - e_0) T_s^4 \\ &+ A_R F_{H,R} e_H (1 - e_X) (1 - e_0) \left(\frac{1}{1 - Z} \right) \sigma T_H^4 \dots [28] \end{aligned}$$

But $A_R F_{0,R} \sigma T_R^4$ = energy radiated from receiver strip past the shield which is equal to that absorbed by the sample, heater case, and surroundings.

Then

$$\begin{aligned} A_R e_R F_{0,R} T_R^4 &= A_R e_R \sigma T_R^4 \left[F_{0,R} e_0 + F_{H,R} (1 - e) \times \right. \\ &\left. (1 - e_X) e_H \left(\frac{1}{1 - Z} \right) + F_{0,R} F_{s,0} (1 - e_0) (e_s) \right] \dots [29] \end{aligned}$$

But the value of $e_s = 1$ when the surroundings are large compared to the measuring system and

$$\begin{aligned} T_R^4 &= T_R^4 \left[e_0 + \frac{F_{H,R}}{F_{0,R}} (1 - e_0) (1 - e_X) e_H \left(\frac{1}{1 - Z} \right) \right. \\ &\left. + F_{s,0} (1 - e_0) \right] \dots [30] \end{aligned}$$

From Equations [29] and [30] with $T_s \cong T_R$

$$\begin{aligned} \frac{K(Mv)}{\sigma F_{0,R}} &= e_0 (T_0^4 - T_R^4) + e_H \frac{F_{H,R}}{F_{0,R}} (1 - e_X) (1 - e_0) \times \\ &\left(\frac{1}{1 - Z} \right) (T_H^4 - T_R^4) \dots [31] \end{aligned}$$

Note that the term for radiation from surrounds cancels the irradiation of the surrounds from the thermopile receiver.

Let

$$Q = e_0 F_{0,R} \sigma (T_0^4 - T_R^4) \dots [32]$$

and

$$\gamma = e_H (1 - e_X) (1 - e_0) \left(\frac{1}{1 - Z} \right) F_{H,R} (T_H^4 - T_R^4) \dots [33]$$

then

$$K(Mv) = Q + \gamma = Q(1 + \gamma/Q) \dots [34]$$

or

$$Q = \frac{K(Mv)}{1 + \gamma/Q} = K(Mv) - \gamma \dots [35]$$

From Equations [32] and [35]

$$e_0 = \frac{K(Mv) - \gamma}{F_{0,R} \sigma (T_0^4 - T_R^4)} \dots [36]$$

Because γ is a function of e_0 , further manipulations will yield an explicit equation for e_0 . From Equation [32]

$$e_0 = \frac{Q}{F_{0,R} \sigma (T_0^4 - T_R^4)} \dots [37]$$

and from Equation [35]

$$e_0 = \left(\frac{1}{1 + \gamma/Q} \right) \frac{K(Mv)}{F_{0,R} \sigma (T_0^4 - T_R^4)} \dots [38]$$

dividing Equation [33] by [32]

$$\gamma/Q = \frac{F_{H,R}}{F_{0,R}} \frac{e_H(1 - e_X)(1 - e_0)}{e_0(1 - Z)} \frac{(T_H^4 - T_R^4)}{(T_0^4 - T_R^4)} \quad [39]$$

Let

$$B = \frac{F_{H,R}}{F_{0,R}} \frac{e_H(1 - e_X)}{(1 - Z)} \quad [40]$$

then

$$C = \frac{T_H^4 - T_R^4}{T_0^4 - T_R^4} \quad [41]$$

$$\gamma/Q = \frac{B(1 - e_0)C}{e_0} \quad [42]$$

Now let

$$e_g = \frac{K(Mv)}{F_{0,R}\sigma(T_0^4 - T_R^4)} \quad [43]$$

which is the emissivity of the surface as computed (see Equation [11]) without corrections for radiation from other surfaces than that from the sample surface. From Equation [38] and [43]

$$e_0 = e_g \left(\frac{1}{1 + \gamma/Q} \right) \quad [44]$$

Substituting Equation [42] into [44] and solving for

$$e_0 = \frac{e_g - BC}{1 - BC} \quad [45]$$

In complete form

$$e_0 = \frac{\frac{K(Mv)}{F_{0,R}\sigma(T_0^4 - T_R^4)} - \frac{F_{H,R}}{F_{0,R}} \frac{e_H(1 - e_X)}{(1 - Z)} \frac{(T_H^4 - T_R^4)}{(T_0^4 - T_R^4)}}{1 - \frac{F_{H,R}}{F_{0,R}} \frac{e_H(1 - e_X)}{(1 - Z)} \frac{(T_H^4 - T_R^4)}{(T_0^4 - T_R^4)}} \quad [46]$$

Equation [44] can be written as

$$\frac{e_0}{e_g} = \left(1 - \frac{\gamma/Q}{1 + \gamma/Q} \right) \quad [47]$$

This expresses the ratio of the correct emissivity of the surface to that emissivity which is computed directly from radiometer readings in Equation [11] and does not account for the addition of energy from the heater case.

From Equation [42]

$$e_0/e_g = 1 - \left[\frac{\left(\frac{1 - e_0}{e_0} \right) BC}{1 + \left(\frac{1 - e_0}{e_0} \right) BC} \right]$$

or

$$\frac{e_0}{e_g} = 1 - \left[\frac{BC}{\left(\frac{e_0}{1 - e_0} \right) + BC} \right] \quad [48]$$

Discussion

J. P. HARTNETT.⁵ The authors have contributed additional

data in a relatively unexplored region of heat transfer. These total emissivity data, however, cannot be directly interpreted as total absorptivities, as stated by the authors, for this equation holds only if the spectral emissivity is independent of wave length and of temperature. As shown by one of the authors⁶ this is not true for some of the surfaces presented in this paper.

The present experimental technique could be improved by placing a black body of the same size and temperature as the sample at the same distance from the radiometer. The radiometer could then be directed at the sample and then at the black body yielding the emissivity directly and eliminating the shape-factor calculation which may be a source of error.

Some comments on Fig. 23 would be appreciated. What happens if the indicated test cycle is repeated?

It is realized by the writer that more recent spectral-radiation data have been obtained by the authors and it is hoped that such information will be made available shortly to workers in the field of heat transfer.

AUTHORS' CLOSURE

(1) Dr. Hartnett is correct in that total absorptivities can only be equal to total emissivities provided the spectral distribution of irradiation is ideal. Actually, ϵ_{irr} is approximately equal to α_{irr} for solar irradiation as used in Equation [1].

(2) Use of an ideal radiator in this experiment would have made the work long and arduous and would really have served no purpose when considering the accuracy desired. A grooved lamp-black plate could have been used with a separate heater and would have required much less time than the use of a holraum (which has been used by various experimenters). Even so, the complexity of maintaining two surfaces at the same temperature is considered unnecessary and superfluous. Our experience in the laboratory with engineering measurements in radiation shows that the technique used with the relatively simple system explained in the text yields data which are sufficiently accurate for general engineering use. A major intention of this paper is to show in as thorough a manner as possible that "good" data of this sort can be obtained with equipment which is available in most laboratories.

A recent experience which we had in helping develop special surfaces for the dissipation of heat by radiation in the Bevatron at the University of California indicated that our equipment supplied accurate data quickly so that the development was finished in a matter of two weeks. Use of cumbersome spectral analytical equipment or special ideal-radiator techniques would have stretched the project (involving as many as 15 different surfaces) to a period of six months.

A lack of knowledge of the utilization of the radiation equation and the general properties of radiation measuring systems leads one to believe that highly precise experiments are the only ones valid. In an experiment using a Perkin-Elmer infrared spectrometer and a high-temperature reflectometer, Dunkle⁷ checked the results against those obtained with equipment similar to that described in this paper and found negligible discrepancy.

The shape-factor computation yields errors well within the dispersion of the data.

(3) A repetition of the data as shown in Fig. 23 with the same plate was not possible. The high temperatures caused a change in condition at the surface which resulted in a reduction of the weathering effect. Magnesium alloys weather rapidly because of the high reactivity of magnesium to water. Heating of the plate apparently reverses this effect to some extent.

⁵ Assistant Professor of Mechanical Engineering, University of Minnesota, Minneapolis, Minn. Assoc. Mem. ASME.

⁶ "Thermal-Radiation Tables and Applications," by R. V. Dunkle, Trans. ASME, vol. 76, 1954, pp. 549-552.

⁷ R. V. Dunkle, private communication

Thermal Conductivity and Its Variability With Temperature and Pressure

By L. S. KOWALCZYK,¹ DETROIT, MICH.

This paper gives a summary of the present status of the theory of thermal conductivity and attempts to explain its variability with temperature and pressure by means of the nature of heat, structure of matter, and resistances offered by matter to heat conduction at various physical states. Experimental data are used to illustrate theoretical considerations; long mathematical derivations are thoroughly avoided.

INTRODUCTION

ALTHOUGH thermal conductivity is a commonly referenced property of materials in engineering work, especially when dealing with heating or cooling, very often its nature is not completely comprehended. This is especially true for work at very low temperatures.

From the physics point of view it is a nonequilibrium property which occurs only in thermally disturbed systems and hence its determination is based on measurements of the time necessary to restore thermal equilibrium.

A definition based on Fourier's law for the conduction of heat states that thermal conductivity represents the amount of heat conducted per unit of time through a unit area (perpendicular to the direction of heat transfer) when the temperature gradient across the heat-conducting element is one unit. In English engineering units, therefore, the dimension of thermal conductivity is²

$$\frac{\text{Btu}}{\text{hr} \times \text{sq ft} \times \frac{\text{deg F}}{\text{ft}}}$$

In scientific works it is either expressed in

$$\frac{\text{cal}}{\text{sec} \times \text{sq cm} \times \frac{\text{deg C}}{\text{cm}}}$$

or in

$$\frac{\text{watts}}{\text{sq cm} \times \frac{\text{deg C}}{\text{cm}}}$$

Numerical values of thermal conductivities of diverse materials at 32 F vary over a large range, i.e., from

¹ Associate Professor, Department of Chemical Engineering, University of Detroit.

² The abbreviated form

$$\frac{\text{Btu}}{\text{hr} \times \text{ft} \times \text{deg F}}$$

is not recommended since it does not express the real character of thermal conductivity.

Contributed by the Heat Transfer Division and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 3, 1954. Paper No. 54-A-90.

$$\frac{241 \text{ Btu}}{\text{hr} \times \text{sq ft} \times \frac{\text{deg F}}{\text{ft}}}$$

for pure silver to about 0.004 and less for some heavy organic gases and vapors. Some illustrative values for the thermal conductivity of various materials at 32 F are shown in Table 1.

TABLE 1 THERMAL CONDUCTIVITIES OF VARIOUS MATERIALS AT 32 F IN BTU/[HR SQ FT (DEG F/FT)]

Silver.....	241
Aluminum.....	132
Zinc.....	65
Iron.....	42
Periclase (MgO).....	24
Quartz.....	11 (parallel to axis)
Corundum (Al ₂ O ₃).....	6
Marble.....	1.6
Ice.....	1.28
Fused quartz.....	1.1
Pyrex glass.....	0.605
Water.....	0.319
Ethyl alcohol.....	0.105
Hydrogen.....	0.101
Benzene.....	0.086
Air.....	0.014
Pentane.....	0.0074
Chloroform (vapor).....	0.0038

Thermal conductivity of various materials changes with temperature in manifold ways; for example, the thermal conductivity of pure metals and nonmetallic crystalline substances increases with the drop in temperature, reaches a maximum in the vicinity of the absolute zero, and then falls to zero.

The highest thermal conductivities thus far measured at these maxima are the following:

For metals

$$\frac{7050 \text{ Btu}}{\text{hr} \times \text{sq ft} \times \frac{\text{deg F}}{\text{ft}}}$$

for a natural copper crystal at 20 K (investigated by Schott, reference 24)³ and

$$\frac{3750 \text{ Btu}}{\text{hr} \times \text{sq ft} \times \frac{\text{deg F}}{\text{ft}}}$$

for 99.996 per cent crystal of aluminum at 14 K (2).

For dielectric crystals

$$\frac{3465 \text{ Btu}}{\text{hr} \times \text{sq ft} \times \frac{\text{deg F}}{\text{ft}}}$$

for a 3-mm-diam single crystal of synthetic sapphire at 50 K (4).

Analysis of thermal conductivity data of various materials at room temperature reveals several facts which throw some light on the nature of this property. In brief, they are the following:

1 Material in a crystalline form conducts heat better than in an amorphous form.

2 In crystals (for instance, in crystalline quartz) and in

³ Numbers in parentheses refer to the Bibliography at the end of the paper.

fibrous materials (wood) thermal conductivity differs with the direction of heat transfer.

3 Small structural differences in crystals due to various conditions of crystal growing also influence their thermal conductivity. For this reason, various types of mica (42) and graphite (40) show different thermal conductivities.

4 Chemical impurities in crystalline substances lower their thermal conductivities considerably. For example, pure metals have much higher thermal conductivities than respective alloys.

5 Solids are better conductors than their respective liquids (for instance, ice conducts heat better than water).

6 A substance in a liquid form shows higher thermal conductivity than in gaseous state.

These facts clearly demonstrate that thermal conductivity is in close relation to the structure and physical state of matter, and all the phenomena connected with thermal conductivity may be explained by a simultaneous application of two theories, i.e., nature of heat and structure of matter.

THE NATURE OF THERMAL CONDUCTIVITY

The present theory of heat is associated closely with the internal energy of matter (52) which in thermodynamics is referred to as the energy connected with the physical and chemical state of a body, i.e., with the configuration and motion of the molecules and atoms within the body. Although incomplete, the dynamic theory of heat permits important conclusions to be drawn with respect to thermal conductivity and its nature. These conclusions, generally confirmed by experiment, may be summarized as follows:

1 Since heat as energy is associated with translational, rotational, and vibrational motions of the molecules, atoms, and their parts, heat transfer by conduction must then be strictly connected with these motions.

2 Temperature increases the intensity and frequency of atomic motions; therefore, thermal conductivity should increase with temperature.⁴

On the other hand, thermal conductivity is closely related to the structure of matter. The closer the contact between molecules, atoms, and their parts within a body, the greater is the possibility of transfer of molecular and atomic motions between neighboring molecules and the higher thermal conductivity.

In accordance with the theories of heat and structure of matter, it is generally accepted today that heat conduction in amorphous solids, liquids, and gases is due to direct transfer of molecular and atomic motions from molecule to molecule (or from atom to atom) at the contact areas. This type of heat transfer is frequently considered as a process of heat diffusion. However, in more organized matter, e.g., in crystals, the atomic motions are converted into a vibratory motion of the whole crystal lattice.

The mechanism of heat transfer by conduction based on the dynamic theory of heat and structure of matter will now be discussed in solids, liquids, and gases. In this discussion emphasis has been placed on physical phenomena and the nature of thermal conductivity and not on mathematical derivations.

SOLIDS

The present theory of heat conduction in solids makes a clear distinction between nonmetallic (dielectric) and metallic sub-

stances. In the dielectric solids heat conductivity depends only on lattice waves produced by atomic motions. In metals and solids conducting electricity, however, two different heat carriers occur—the lattice waves and the electrons. Detailed, mathematical treatments of the theory of heat conduction in solids were given by Makinson (30), Sondheimer (49), Klemens (28), de Boer (7), and others.

NONMETALLIC (DIELECTRIC) SOLIDS

Thermal conductivity in nonmetallic solids is entirely due to the atomic motions and their transfer from hot to cold regions. At present, it is generally accepted that these motions assume the form of thermal vibrations which in turn can be regarded as a superposition of thermoelastic (acoustic) waves known as lattice waves.

The concept of the lattice waves as a heat-conducting medium in solids and the concept of thermal resistance due to scattering of the lattice waves by themselves and other obstacles on their path were first introduced by Debye (14) who was able to explain the difference between the thermal conductivities of dielectric crystals and amorphous solids. According to this theory, each atom in a solid has a fixed equilibrium position, and its thermal motion can be resolved into normal modes of vibration which are traveling elastic (sound) waves.

Debye derived a theoretical formula for the thermal conductivity of solids as a simple product of four terms

$$k = \frac{1}{4} \lambda u \rho C_m \dots \dots \dots [1]$$

where

λ = average distance of effect of thermoelastic wave, understood as distance which wave travels before its intensity is reduced to $1/e$ of its initial value

u = average velocity of propagation of waves

ρ = density of solid

C_m = specific heat per unit of mass

Peierls (36) further developed the Debye theory. Debye treated the atomic vibrations of solids as a collection of thermoelastic waves in a continuum. Peierls considered the individual atomic motions and quantized the lattice waves into quanta of vibration adopting the concept of a phonon (21) analogous to the concept of a photon in the radiation theory. (A phonon is equivalent to one quantum of energy in the form of a thermoelastic wave of a fixed frequency.) In other words, Peierls discontinued the lattice waves, and assimilated heat conduction in solids to that of gases replacing the lattice waves by so-called "phonon gas." Consequently, the modified Debye theory leads to the same type of formula for the thermal conductivity of solids as derived for gases from the kinetic theory

$$k = 1/3 l u C_v \dots \dots \dots [2]$$

where l is the mean free path, defined in the same manner as λ in Equation [1]; u is the velocity of the lattice waves (phonons); and C_v is the specific heat per unit volume.

Since Equation [2] is frequently used to interpret thermal conductivity of solids, let us discuss it briefly. The meanings of u and C_v terms are obvious; l is the distance which the lattice waves (phonons) travel before their intensity is attenuated to $1/e$ of their original value. Therefore the length of l depends upon the scattering of the lattice waves (phonons) by various processes and may be considered as a parameter characteristic for scattering processes. It also may be referred to as the effective length of relaxation of the lattice waves to $1/e$ by various scattering processes.

⁴ There are two exceptions to this rule:

(a) Thermal conductivity of dielectric crystals and metals decreases from a maximum value at very low temperatures toward the melting point.

(b) Thermal conductivity of the majority of liquids decreases with temperature.

These deviations are discussed in detail in subsequent parts of this paper.

According to the Debye theory, scattering of the lattice waves gives rise to thermal resistance. The thermal resistance is the reciprocal of the thermal conductivity, or

$$R = \frac{1}{k} = \frac{3}{luC_v} \dots \dots \dots [3]$$

where R is the thermal resistance. Equation [3] shows that the thermal resistance is inversely proportional to the mean free path or, in other words, strong scattering processes diminish l and increase thermal resistivity.

(a) *Ideal Crystals.* Since in ideal crystals the displacement of an atom from its equilibrium position is proportional to the force, the atomic vibrations in an ideal lattice are essentially harmonic and any number of the lattice waves can be superimposed without mutual interference. For this reason the lattice waves (phonons) in ideal crystals are not scattered; each wave travels independently of the others and encounters no resistance on its way. In terms of Equation [2], l is very large and, therefore, the ideal crystals should have an infinite thermal conductivity. Also, the sound velocity in the ideal crystals (transmitted by independent waves) is very high. For instance, for quartz crystals it would be from 550,000 to 770,000 cm/sec (24). This is another factor increasing the magnitude of the thermal conductivity in the ideal crystals.

Finally, the normal modes obtained from resolving of lattice vibrations in the ideal crystals turn out to be plane traveling waves.

(b) *Actual Crystals.* In actual crystals, however, the force on an atom is not proportional to the displacement from its equilibrium position. This fact is chiefly responsible for the anharmonicity of the atomic vibrations in actual crystals or for mutual scattering of the lattice waves (collisions between the phonons). In large, perfectly formed crystals, free from chemical impurities and mechanical damages or strains, this is the only cause of scattering of the lattice waves. In actual crystalline substances, however, there occur other factors, such as specimen boundary, crystallographic imperfections of the lattice, mechanical defects of the lattice (cracks, strains, etc.), chemical impurities, etc., which also scatter and distort the lattice waves, lessen their intensity and reduce their velocity.⁵

Scattering processes lead to the reduction of l and u in Equation [2] and consequently to the diminution of heat conductivity or to the increase of thermal resistance, Equation [3]. The more the lattice waves are scattered, the greater is the resistance offered to their transmission.

In actual crystals we can distinguish four groups of scattering processes:

Mutual scattering of the lattice waves (the three-phonon processes).

Elastic scattering.

Inelastic scattering.

Higher-order phonon interaction (Pomeranchuk, 1941).

These processes will now be discussed.

Mutual scattering of the lattice waves arises from anharmonicities of the interatomic forces. According to the Peierls theory of phonon-gas, there are two classes of processes causing the mutual scattering of the "phononized" lattice vibrations (5). The theory concludes that such interaction between the phonons leads either to the destruction of one phonon and creation of two new ones, or vice versa. For this reason these perturbations are commonly referred to as the three phonon processes. In the first class of these, the total energy and direction of the energy current

remain unchanged. In other words, this type of the phonon collisions does not contribute directly to the increase of heat resistance. Its influence, rather, is indirect since it might change the state of polarization of the phonons and render them more susceptible to the second class of collisions in which the total energy of colliding phonons is conserved but direction of flow of energy is altered. Since this class of collisions introduces impedance to the flow of energy, it must give direct rise to thermal resistance.

Peierls has termed this second class of phonon collisions the "Umklapp" processes (very frequently referred to as the "U" processes), the best translation of which is the "flapping-over" processes between two phonons. In order to be liable to a U-collision, the phonons must possess a certain minimum of energy determined by Peierls as $k\theta/2$ where k is the Stefan-Boltzmann constant and θ is the Debye specific-heat parameter (the Debye temperature). The Umklapp processes are the most important components of thermal resistance in pure, regular crystals and are therefore mainly responsible for heat conductivity and establishing temperature equilibrium.

Elastic and inelastic scatterings arise from defects in the lattice, boundaries, microcrystalline structure, etc., and are often termed "structure scattering." During elastic scattering the energy of phonons remains unchanged.

It has been proved experimentally that irregularities in the lattice of atom dimensions (small-scale lattice defects) scatter phonons and increase thermal resistance. Berman (4) measured the effect of displaced atoms produced in a quartz crystal by neutral irradiation. The results indicated that at low temperatures the thermal conductivity of the quartz crystal distinctly decreased with the concentration of displaced atoms. Estermann and Zimmerman (18) also showed that between 10 and 20 K, the thermal conductivity of the cold-worked specimens of 90-10 Cu-Ni alloy was only about one third of that of the single crystal. This fact indicated that in the cold-worked specimens stronger scattering of the phonons occurred due to "dislocations and short-range lattice defects resulting from cold-working." On the other hand, annealing increased the thermal conductivity; i.e., it removed to a certain degree the deformation of the lattice.

Undoubtedly, larger defects in the crystal lattice (mechanical breaks and cracks, crystallographic defects, etc.) will contribute to further scattering of the phonons.

The lattice waves (phonons) also are scattered at the boundaries of a specimen and at the grain boundaries (crystal walls) of polycrystalline materials. In other words, the larger the crystal, the smaller is the scattering at its walls and the higher its thermal conductivity. This explains why single crystals show higher thermal conductivity than polycrystalline structures of the same material. This is especially true at very low temperatures. Experiments of de Haas and Biermasz (23) on four different sizes of KCl crystals in temperatures as low as 2 K and those of Berman (4) on two different-size single crystals of synthetic sapphire between 2 and 200 K revealed that in temperatures below the maximum of thermal conductivity (7 K for KCl and about 50 K for sapphire), the influence of boundary scattering (grain size) on the thermal conductivity of the samples was very pronounced and larger specimens showed higher thermal conductivities.

The experimental results of de Haas and Biermasz gave a basis to Casimir (12) for the theoretical treatment of this problem. He proved that at sufficiently low temperatures, the mean free path l in the modified Debye formula, Equation [2], is proportional to the smallest linear dimension of the crystal.

Probably the most important process of phonon scattering in this group of processes is scattering by chemical impurities. It is a well-known fact that the thermal conductivity of pure metals is

⁵ According to Jakob (24), the velocity of heat conduction in quartz is about 0.02 cm/sec, or 30-40 million times smaller than that of a sound wave.

very sensitive on impurities. For example, the slightest trace of arsenic in copper reduces materially its thermal conductivity (to about $1/3$ of that for pure copper). Chemical impurities introduce foreign atoms into the crystal lattice and these in turn scatter phonons (and electrons in metals). Berman (4) measured the resistance caused by chromium impurities in artificial sapphire between 5 and 100 K. He found that the impure crystal (ruby) has lower thermal conductivity than sapphire although in this particular case the scattering by impurity is quite small.

Phonon scattering of a higher order is caused by mosaic structure of a crystal which results from a division of the crystal into regions of slightly different orientation. It is another departure from complete regularity in crystal lattice and may result in scattering of phonons. This process, however, is not important at low temperatures, but only at the highest temperatures (39).

Since particular scattering processes are important only in certain ranges of temperature, and vary with temperature in various degrees, and since their contribution to the total thermal resistance is not exactly known, they cannot be used at present for calculation of thermal resistances or conductivities in actual crystals. For example, Berman (4) concluded that for single crystals of sapphire and diamond, the thermal conductivity at the low temperatures cannot be entirely explained by the U-processes and boundary scattering and an additional resistance independent of temperature must be assumed. Some authors have postulated the presence of clusters of defects resulting in this additional resistance.

The known scattering processes, however, reveal a general character of the thermal resistance-temperature and thermal conductivity-temperature curves.

(c) *Glassy (Amorphous) Materials.* Glassy or amorphous materials are solids in which molecules (or atoms) are at random; i.e., their structure is lacking both symmetry and periodicity. The x-ray analysis reveals, however, that the arrangement of atoms in glassy materials shows short-range order (similar to that in liquids), but no long-range order (as in crystalline substances). Each atom in a glassy material has a fixed equilibrium position about which it oscillates. The atomic thermal motions are then resolvable into normal modes of vibration but these modes do not combine into the lattice waves and the thermal vibrations of each atom interact with those of the neighboring atoms and interchange energy (inelastic scattering). Irregularity in the glass structure and its direct consequence, inelastic structure scattering of atomic vibrations, result in low thermal conductivity of glasses. For instance, experiments of Berman (4) showed a distinct difference between the thermal conductivities of quartz crystal and quartz glass. At 50 K, the thermal conductivity of quartz crystal was about 170 times larger than that of quartz glass, whereas at 10 K (the maximum value for the quartz crystal) the ratio increased to 10,000. At 0 C this ratio is about 10.

Distinct differences in thermal conductivities also have been observed for diamond, graphite, and amorphous carbon, for crystalline and amorphous sulphur, etc.

(d) *Refractory Materials.* Refractory materials may be considered as mixtures of crystals and amorphous materials. Their thermal conductivity, therefore, may be assumed to be the sum of contributions of crystalline and glassy components; i.e., it will depend upon the thermal conductivities and volume fractions of the constituents.

(e) *Mixtures of Amorphous Materials.* The remaining solids represent heterogeneous materials of various composition and structure. To this group belong building and insulating materials, plastics, rubber, leather, fibrous and granular materials, powders, etc. In a broad sense, they are mixtures of solid amorphous materials and amorphous solids with fluids. Their thermal conductivity (or rather, apparent thermal conductivity) will also depend

upon the thermal conductivities and volume fractions of the components. If one of the components is a fluid (for example, air in the pores of an insulating material), heat transfer by conduction and radiation (only in gases) also should be taken into account (1).

METALS

Thermal conductivity of metals and semiconductors (graphite, metallic oxides, sulphides, silicates, etc.) depends upon two separate mechanisms:

- 1 Heat conduction by the lattice waves (the phonon-gas).
- 2 Heat conduction by free electrons (the "electron-gas").

Thus the total heat conductivity of a metal appears to be the sum of two terms, the lattice conductivity k_l and electronic conductivity k_e ,

$$k = k_l + k_e \dots \dots \dots [4]$$

In pure metals of high electrical conductivity, electronic thermal conductivity is a dominating factor and may, in some cases, be more than 100 times higher than the lattice conductivity. In alloys, however, both types of conductivity may be of the same order of magnitude.

Scattering of the lattice waves by various processes discussed in the preceding paragraph originated the concept of thermal lattice resistance. Similarly, the concept of electronic thermal resistance is based on various processes which scatter the stream of free electrons. There are two processes which mainly contribute to the electronic thermal resistance—scattering of electrons by thermal vibrations of the lattice and by impurities and small-scale lattice defects. Obviously, the lattice waves will be simultaneously scattered by electrons. In metals, therefore, there will be an additional term in the lattice thermal resistance, i.e., scattering of phonons by electrons.

A mathematical treatment of the electronic contribution to the thermal conductivity of metals is given by Sondheimer (49), de Boer (7), and others.

(a) *Determination of k_l and k_e in Metals.* There are several methods used for the determination of the lattice conductivity in conductors of electricity independently of the electronic conductivity. The most common are the following:

- 1 At room temperature and above, the thermal and electrical conductivities of pure metals are related by the Wiedemann-Franz law

$$k = \sigma LoT \dots \dots \dots [5]$$

where

k is the thermal conductivity, σ is the electrical conductivity, Lo is the Lorentz constant [2.45×10^{-9} (watt-ohm/deg K²)], and T is the absolute temperature.

Assuming similar mechanism for the transportation of heat and electricity by free electrons, one may conclude that only

$$k_e = \sigma LoT \dots \dots \dots [5a]$$

and calculate the lattice conductivity from Equation [4] by subtracting k_e from k total, determined experimentally

$$k_l = k - \sigma LoT \dots \dots \dots [6]$$

This method, however, cannot be used below room temperature since the Wiedemann-Franz law is not valid at low temperatures.

- 2 A strong magnetic field reduces considerably the electrical conductivity of metals and then their electronic thermal conductivity, Equation [5a]. The thermal conductivity of metals, determined in a magnetic field, therefore, will be only the lattice conductivity k_l .

3 Thermal conductivity of crystals depends upon the grain size and at constant temperature may be approximately represented by an equation

$$k = \frac{a}{b + n} \dots \dots \dots [7]$$

where n is the average number of grain boundaries per unit length in the metal in the direction of heat flow and a and b are constants. Since k is not affected by grain size, then Equation [7] relates only to k , which may be determined by measurements of total k in several samples of the same material differing only in grain size.

(b) *Alloys. The structure of alloys.* Alloys are mixtures of two or more metals and in the solid state are composed of crystals but their structure depends upon the composition and properties of the constituents. Classification of alloys is based on their behavior at freezing. From many possible combinations, only four of the most common binary-alloy structures will be discussed.

1 Continuous solid-solution alloys. In this class, when the melt is cooled to freezing temperature, it forms a continuous series of homogeneous solid solutions. This class is represented by silver-gold and copper-nickel binary alloys.

2 Eutectic alloys. The components are soluble in the liquid phase but immiscible in the solid phase. At freezing they form either a mechanical mixture of two pure metal crystals (at a eutectic composition) or deposit crystals of one metal until the eutectic composition is reached. Typical alloys in this class are cadmium-zinc, silver-lead, lead-antimony, and others.

3 Limited solid solutions. The metals are completely miscible in the liquid state but only partially soluble in each other in the solid state. The lead-tin alloy is characteristic of this class.

4 Intermetallic compounds. Alloying metals at freezing may form intermetallic compounds in definite stoichiometric ratios. For example, for lead-magnesium, such a compound is $PbMg_2$. Intermetallic compounds show an ordered homogeneous structure and behave like single components.

Thermal Conductivity of Alloys. The thermal conductivity of alloys, similarly to that of pure metals, is the sum of two terms, the lattice conductivity and the electronic conductivity. For this reason it should be considered from the point of view of phonon and electron scattering processes.* Since these processes depend upon the structure of alloys, each class will be discussed separately.

1 The variation of thermal conductivity in continuous solid-

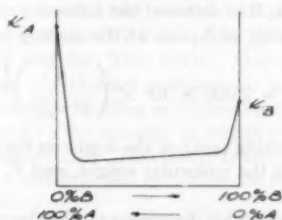


FIG. 1 VARIATION OF k WITH COMPOSITION IN CONTINUOUS SOLID-SOLUTION ALLOYS

solution alloys with their composition is shown in Fig. 1 where A and B are two pure metals.

In this class of alloys, an atom of the alloying metal may be substituted for a solvent metal atom at random points in the solvent lattice. Also, it may be located in the interstices of

the solvent lattice (the interstitial type of solution). Even a very small addition of alloying metal (or introduction of foreign atoms into the solvent lattice) rapidly increases scattering of phonons and electrons and reduces both the lattice and electronic conductivities. Apparently, the lattice conductivity suffers more than the electronic conductivity since the propagation of the lattice waves requires a perfect order in a crystalline structure. Small amounts of foreign atoms in the lattice weaken the lattice conductivity which sharply drops (extreme parts of the plot in Fig. 1). When the amount of alloying metal is large enough, irregularities in the lattice structure suppress completely the lattice waves and the lattice conductivity becomes somewhat similar to that of amorphous materials (inelastic scattering of thermal vibrations). Further additions of alloying metal have a negligible influence on the structure of the alloy and the thermal conductivity remains independent of composition (a horizontal part of the plot in Fig. 1).

The greater the difference between the atomic structures of the solvent metal and the impurity, the higher the drop will be in thermal conductivity. For this reason, metalloids cause a much larger drop of k in metals than other metallic impurities.

2 In eutectic alloys (alloys with complete mutual insolubility of components in the solid phase) which consist of a mechanical mixture of two types of crystals, the change of thermal conductivity with composition may be roughly represented by a straight line connecting the thermal conductivities of pure components. (Actually, this is a curve falling below the straight line. It is due to scattering action of crystal walls in a heterogeneous mixture.) This fact indicates that pure crystalline structure of both components is preserved at any composition and heat is conducted independently by both types of the crystals. Thermal conductivity of eutectic alloys is, therefore, the sum of contributions of both components or an additive property (as in refractory materials).

3 In limited solid solutions, the thermal conductivity-composition curve assumes a shape according to the behavior of the solid phase; k decreases rapidly with increasing alloy content in the continuous solid-solution parts of the alloy. However, when the limit of solid solubility is reached and two different crystals are formed, the thermal conductivity varies approximately linearly as for eutectic alloys.

4 If the alloying metal forms intermetallic compounds with the solvent metal, their existence is marked on the thermal conductivity-composition curve by maxima. Such a component has a definite crystalline structure and undoubtedly offers smaller resistance (less scattering) to the flow of phonons and electrons than a continuous solid-solution alloy. Maxima in thermal conductivity of solid-solution alloys were also observed at certain compositions which do not justify the existence of intermetallic compounds (3). They may be explained, however, by an ordered structure of an alloy (so-called superstructure or superlattice) appearing after certain heat-treatments. The main difference between an intermetallic compound and an ordered alloy structure is that the compound is stable up to its melting point whereas the superstructure may be destroyed before, for example, by quenching.

Many attempts have been made to relate thermal conductivity of alloys (particularly that of various kinds of iron and steel) to their composition. Since in steel several impurities are present, the problem becomes more complex. In general, their influence on k of pure metal is roughly additive.

(c) *Semiconductors.* In this class of materials, heat conduction by phonons is supported by electronic conduction. The number of free electrons per unit volume in semiconductors, however, is usually small and their contribution to total thermal conductivity depends upon temperature. In low temperatures the

* A detailed discussion on the influence of impurities on the electronic thermal conductivity is given by Sondheimer (49).

electronic thermal conductivity is insignificant. However, it rapidly increases with temperature and at high temperatures it may even outweigh the lattice conductivity.

THERMAL CONDUCTIVITY AT A TRANSITION POINT

Under certain conditions, some crystalline polymorphous solids demonstrate a transition point above and below which only one crystalline form is stable. Such transition from one crystallographic form to another is accompanied by a rearrangement of atoms in the lattice and consequently by a change in density and thermal conductivity.

Experiments of Kaye and Higgins (25) on sulphur revealed a sudden drop in thermal conductivity when rhombic sulphur (sp gr 2.07, stable below 95.5°C) is converted into monoclinic (sp gr 1.96, stable above 95.5°C). Their results are shown in Fig. 2 (recalculated into English system of units).

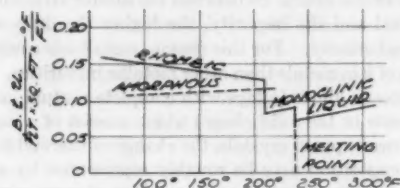


FIG. 2 THERMAL CONDUCTIVITY OF SULPHUR

On the other hand, the transition point from alpha to gamma iron is not marked by a sudden change in thermal conductivity but only by a change in the slope of the temperature-conductivity curve (3). This is probably due to the fact that the thermal conductivity of metals depends chiefly on the motion of electrons which is not much influenced by the transition of the lattice.

LIQUIDS

The Structure of Liquids. The process of melting may be considered as a change of more orderly state of molecules in solids into less organized state in liquids (hence increase in entropy). This so-called "order-disorder" transformation introduces important changes into the molecular structure; it loosens the bonds among the molecules, destroys the order of the solid structure, and creates a possibility of free (thermal) movement of the molecules.

The readiness with which liquids flow and diffuse led in the past to a general belief that the structure of liquids is close to that of gases and the molecules in liquids are at complete randomness. It was assumed that the only difference between liquids and gases lies in the larger average distance and, therefore, in the smaller attractive forces among molecules.

However, in the vicinity of the melting point the density of a liquid is not greatly different from that of a solid and there cannot be much possibility of free movement of the molecules within liquids. For this reason the modern theory of the structure of liquids assumes that structurally liquids are closer to solids than to gases and retain the "short-range" order (like glasses) to a great extent. This point of view is supported by x-ray diffraction measurements of liquids and the existence of so-called "liquid crystals," which definitely exhibit some properties of anisotropic crystals.

Thermal Conductivity of Liquids. Changes in the structure of matter at melting have a pronounced influence on the thermal conductivity of liquids which abruptly decreases at the melting point. This phenomenon can be explained either by the destruction of more or less organized solid structure (excluding heat propagation by the lattice waves as in crystals) or by looser com-

pactness of molecules in the liquid state and, therefore, by greater amplitude of the atomic vibrations.⁷

Experimental data on thermal conductivity of solid and liquid phases in the vicinity of the melting point are available for helium, water, sulphur, and many metals. Except for bismuth, all of them indicate a sudden drop of k at melting point as shown for monoclinic sulphur in Fig. 2 (data of Kaye and Higgins). On the same figure, experimental data of Jakob (24) for amorphous sulphur (containing a little of crystalline sulphur) are represented by a dashed line.

Thermal conductivities of bismuth in solid and liquid phases are given in Fig. 3.

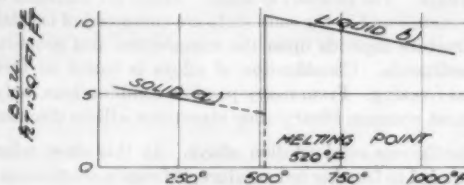


FIG. 3 THERMAL CONDUCTIVITY OF BISMUTH IN SOLID AND LIQUID PHASES

(a) J. H. Perry (38), p. 456. (b) R. C. L. Bosworth, (8), p. 84.]

The only explanation of this phenomenon seems to lie in an increase of the lattice contribution (atomic motions) since it is found that the velocity of sound in liquid bismuth is greater than that in solid bismuth.

Many authors have attempted to derive an expression relating thermal conductivity of liquids with other properties which are easier to measure. In the following discussion these attempts are divided into two groups. The expressions for k in the first group comprise various properties of fluids except the velocity of sound. In the second group, the acoustic velocity in fluids was a principal property, correlation of which to k was attempted.

One of the earliest equations in the first group is that proposed by Weber (51). A series of empirical equations dimensionally inconsistent was suggested by Smith (48). According to Rao (44), the theory of liquids should be approached from the point of view of the solid state, the density of which is closer to that of the liquid. Rao considered a liquid near its melting point as an assemblage of a large number of harmonic oscillators (atoms) vibrating with an amplitude comparable with the mean molecular distance. Assuming that transfer of energy takes place at each extreme vibration, Rao deduced the following expression for the thermal conductivity of liquids at the melting point

$$k = 2.096 \times 10^6 \times \left(\frac{\theta_s}{MV_s^{1/3}} \right)^{1/2} \dots \dots \dots [8]$$

where θ_s is the melting point of the liquid on the absolute scale of temperature, M is the molecular weight, and V_s is the molecular volume.

Fournier (22) based his discussion on the thermal conductivity of liquids on a cage made up by neighboring molecules from which a molecule can escape only with difficulty. He suggested that the molecules of the liquid in their cages undergo motions intermediate

⁷ It is generally assumed that in stationary fluids heat conductivity is due to the transfer of energy (molecular and atomic motions) by direct contact between molecules which are of different size and shape in various fluids. Obviously, the possibility of free thermal movements of molecules or their groups introduces another method of heat transfer within fluids (heat convection) which is completely independent of heat conduction. In measurements of thermal conductivity of fluids, the influence of heat transfer by convection should be eliminated as it increases the rate of heat transmission.

between the harmonic motion of crystalline solids and the translational motion of gases. On this basis Fournier derived the following formula

$$k = \frac{C_v L^2 T}{V^2 \mu} \quad [9]$$

in which C_v is the molal heat capacity at constant volume, V is the molal volume, L is the mean distance between molecules, T is the absolute temperature, and μ is the coefficient of viscosity.

Denbigh (15) related log of the Prandtl number of a liquid to its absolute temperature and the latent heat of vaporization at constant pressure. Denbigh's equation may be used for estimation of k of a liquid when its viscosity, specific heat, and latent heat of vaporization are known.

Finally, in this group the papers of Palmer (34) and Riedel (46) should be mentioned. Palmer modified Weber's equation introducing in it the Trouton's constant and the molal entropy of vaporization. He also discussed the influence of hydrogen bonds on the thermal conductivity of liquids. Riedel worked out an equation relating thermal conductivity of liquids to reduced temperature.

The early attempts to derive an expression for the thermal conductivity of liquids which implies a relationship between k and the velocity of sound in a respective liquid are those of Brillouin (11) and Paschki (35). Independently of them, Bridgman (9) derived on the same basis the following equation

$$k = \frac{2Ru}{z^3} \quad [10]$$

where R is the universal gas constant in heat units, u is the velocity of sound in the liquid, and z is the mean distance between the centers of molecules in cubical arrangement

$$z = \left(\frac{M}{\rho} \right)^{1/3} \text{ cm}$$

where M is the molecular weight in grams.

Equation [10] gives relatively close results to those obtained by experiment although in some cases deviations are as large as 38 per cent.

Other theoretical equations for k of liquids expressing the thermal conductivity in terms of the velocity of sound were developed by Kincaid and Eyring (26) and Powell, Roseveare, and Eyring (43).

Examination of experimental data on the thermal conductivity of nonmetallic liquids shows that water and liquid ammonia are the best heat conductors—at 32 F, k in Btu/[hr sq ft (deg F/ft)] for water is 0.319 and for NH_3 , 0.312. Organic liquids of a large molecule have the thermal conductivity at room temperature about 0.1 although the drop in thermal conductivity is not strictly proportional to the number of atoms even in the same class of chemical compounds. This fact indicates that an arrangement of atoms in a molecule also influences the transfer of energy between molecules.

There are no experimental data for the thermal conductivity of liquid crystals which having higher structural order should reveal a higher thermal conductivity than ordinary liquids. A sudden change of k is also to be expected at transition points (liquid crystal-normal liquid).

Thermal Conductivity of Molten Metals. Molten metals and electrolytes represent a separate class of liquids since their thermal conduction is based on two different processes, atomic motions and drifts of free electrons. The electronic conduction greatly magnifies the thermal conductivity of these liquids since it is 10 to 1000 times higher than that of nonmetallic liquids. For

instance, k in Btu/[hr sq ft (deg F/ft)] for several molten metals is the following: Sodium at 100 C: 49.13, cadmium at 349 C: 25.8, bismuth at 300 C: 10.16, and mercury at 0 C: 4.74.

The increasing application of liquid metals as heating media created interest in their thermal conductivity and its correlation to other properties of molten metals. Bidwell (6) suggested a formula for k of liquid metals which takes into account the contribution of atomic motions and free electrons; Rao (45) showed that the ratio of thermal conductivities in the solid and liquid

phases near the melting point is given by $e^{\frac{2}{3} \frac{\lambda_m}{k_0 T_m}}$, where k_0 is the Stefan-Boltzmann constant, and λ_m/T_m is the entropy of melting.

GASES

The Gaseous State. The process of evaporation loosens almost completely intermolecular bonds which existed in the liquid state and increases the distance between molecules to such degree that they are free to move in any direction (the volume of saturated water vapor at 212 F and 1 atm is 1604 times larger than that of liquid water at the same conditions) and the only obstacles on their path are other molecules with which they collide. In other words, the gaseous state has neither type of order and the molecules are distributed at a complete randomness. All the properties and behavior of gases are explained in terms of the kinetic theory of gases which is generally adopted.

Heat Conductivity of Gases. The heat conductivity of gases is closely related to the movement of molecules and their collisions during which the exchange of energy takes place. In other words, the heat conduction in gases may be compared to a process of molecular diffusion from hotter to colder regions but restricted to the exchange of location and energy of molecules. Heat transmission in gases by thermal currents (convection) and by electromagnetic waves (radiation) are two different methods of heat transfer and are not discussed here.

A derivation of an equation for the thermal conductivity of gases is based on the kinetic theory of gases and leads to a general expression

$$k = \frac{1}{3} l u C_v \quad [11]$$

in which l is the mean free path (mean distance traveled by the molecules between collisions), u is the mean velocity of the molecules, C_v is the specific heat at constant volume.

The identical expression, Equation [2], has been derived for dielectric crystals assuming in accordance with Peierls' theory that atomic thermal vibrations (the lattice conductivity) can be resolved into streams of phonons (the phonon-gas).

On the other hand, the kinetic theory of gases leads to the following expression for the coefficient of viscosity μ

$$\mu = \frac{1}{3} l u \quad [12]$$

and, in turn, to a new equation for the thermal conductivity of gases. Combining Equations [11] and [12]

$$k = \mu C_v \quad [13]$$

Actual measurements of the thermal conductivity of gases indicate, however, that k is greater than calculated from Equation [13] and should be expressed by a corrected equation

$$k = a \mu C_v \quad [14]$$

where a is a magnifying factor whose numerical values depend upon the nature of a gas. For monatomic gases (of spherical molecules) a has the highest value about 2.5. For diatomic gases

α is somewhat smaller than 2 and is gradually decreasing with the number of atoms in a molecule down to 1 (24). Although a complete theory on the role of factor α in the thermal conductivity of gases has not yet been developed it is believed that it is connected with the rate of exchange of various forms of molecular energies (those of translation, rotation, and intramolecular oscillation) during the collisions of molecules. Eucken (19) supposes that these forms of energies are independent of each other and are exchanged between molecules with diverse degrees of easiness. The faster molecules of monatomic gases are charged with more energy of translation and carry it faster. Since this energy is relatively easily exchangeable during collisions, it results in higher thermal conductivity of monatomic gases for which $\alpha = 2.5$ and is in agreement with theory and experiment.

The energy of translation decreases with the number of atoms in a molecule (slower molecules) but energies of rotation and oscillation will increase. Since they are more difficult to interchange during the molecular collisions, the factor α and the thermal conductivity decrease with the number of atoms in a molecule. In other words, the thermal conductivity of gases is affected through factor α by the size and shape of the molecule as well as by the character of atoms in the molecule.

McAdams (31) gives Eucken's empirical formula for α in terms of the ratio of specific heats γ [$\alpha = 0.25(9\gamma - 5)$], as suggested by Eucken in 1913, whereas Jakob (24) discusses a derivation of another Eucken's formula for α (1940) based on a consideration of three various forms of energy possessed by a molecule. Jakob gives also a procedure recommended by Eucken (1940) for the determination of the thermal conductivity of gases in the absence of reliable experimental data.

Numerical values of the thermal conductivity of gases are scattered over quite a large range as can be seen from Table 2.

TABLE 2 THERMAL CONDUCTIVITY OF SOME GASES IN BTU/HR SQ FT (DEG F/FT) AT 32 F AND 1 ATM

Hydrogen.....	1.101
Helium.....	0.082
Neon.....	0.027
Methane.....	0.0175
Oxygen.....	0.0142
Nitrogen.....	0.0140
Carbon dioxide.....	0.0085
Butane (n-).....	0.0078
Methyl chloride.....	0.0053
Chloroform.....	0.0038

VARIABILITY OF THERMAL CONDUCTIVITY WITH TEMPERATURE

Solids: Nonmetallic (Dielectric) Solids

(a) *Crystals.* Thermal conductivity of dielectric crystals is due to the lattice vibrations "emitted" in the forms of the lattice waves (phonons) which are hampered by various scattering processes; this was discussed in detail in the preceding paragraphs. In general, the phonon scattering processes shorten the mean free path of phonons between collisions, Equation [2], or increase the resistance on the path of the phonons. In the first approximation, these thermal resistances resulting from various scattering processes may be considered as additive and we may write

Total thermal resistance, R_t

= resistance due to mutual scattering of phonons (umklapp-processes), R_u
 + resistance due to boundary scattering, R_B
 + resistance due to defects of lattice, R_D
 + resistance due to chemical impurities, R_I
 + resistance due to higher-order scattering (mosaic structure) R_M

Algebraically

$$R_t = R_u + R_B + R_D + R_I + R_M \dots [15]$$

These resistances depend upon absolute temperature in various degrees. For low temperatures at which only the small wave numbers are important, de Boer (7) gives the following relationships between various causes of resistance and absolute temperature:

Mutual scattering of phonons (U-processes)

$$R_u = \frac{1}{AT} e^{-\frac{1}{2} \frac{\theta}{T}} \dots [16]$$

where A is a numerical constant, θ is the Debye temperature, and T is the absolute temperature.

Boundary scattering

$$R_B = \frac{1}{BLT^2} \dots [17]$$

where B is a numerical constant and L is the smallest linear dimension of the sample.

Scattering by chemical impurities

$$R_I = CT \dots [18]$$

where C is a numerical constant.

Scattering by mosaic structure for small wave numbers (Pomeranchuk, 1942)

$$R_M = \frac{1}{DT} \dots [19]$$

where D is a numerical constant.

In well-formed crystals, free from chemical impurities and lattice and mechanical defects, the thermal resistance is caused by the U-processes and the boundary scattering only. Algebraically

$$R_t = R_u + R_B = \frac{1}{AT} e^{-\frac{1}{2} \frac{\theta}{T}} + \frac{1}{B'LT^2} \dots [20]$$

In this equation $B' < B$, since the relative contribution of boundary scattering to the total resistance is much larger than would result from the addition of these two resistances (Klemens, 1951).

Since the thermal resistance is the reciprocal of the thermal conductivity, Equation [3], we may write

$$k_t = \frac{1}{R_t} = \frac{1}{R_u + R_B} = \frac{1}{\frac{1}{AT} e^{-\frac{1}{2} \frac{\theta}{T}} + \frac{1}{B'LT^2}} \dots [21]$$

This is a basic equation for the thermal conductivity of regular dielectric crystals as a function of absolute temperature—at very low temperatures. The thermal conductivity-temperature curve, based on Equation [21], is shown in Fig. 4.

This purely theoretical curve reveals a maximum of k_t in the vicinity of the absolute zero the exact position of which depends upon the numerical constants involved in Equation [21] and temperature θ . From the maximum the curve falls gradually toward the melting point. At high temperatures (reasonably higher than θ), however, the lattice waves become shorter and scattering of phonons by U-processes becomes more significant. The thermal resistance due to these processes gradually shifts from the exponential form, Equation [16], to a simple T -law (37), or

$$R_u = A'T \dots [22]$$

Since, at higher temperatures, the contribution of boundary scattering becomes negligible, Equation [21] simplifies to

$$k_t = \frac{1}{A'T} \dots [23]$$

as shown by a dashed line in Fig. 4. On the other hand, Pomeranchuk (1942) concluded that at higher temperatures k_i of dielectric solids is proportional to $T^{-1/2}$.

Also, at very high temperatures, $R_u > R_D$ or the thermal conductivity is less sensitive to defects of the lattice.

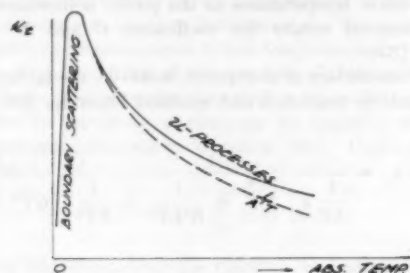


FIG. 4 THERMAL CONDUCTIVITY OF DIELECTRIC CRYSTALS VERSUS TEMPERATURE

All experimental data on the thermal conductivity of dielectric crystals which are available for many materials such as ice, helium, quartz, micas, sapphire, silicon carbide, etc., show a complete agreement with the curve shown in Fig. 4. The numerical value of k_i at the maximum is many times higher than that at ordinary temperatures. For instance, the thermal conductivity of a quartz crystal at the maximum—about 10 K (measured by Berman) was about 580 whereas at 0°C it is 11 (parallel to axis) and 6 (normal to axis).

(b) *Amorphous Materials (Glasses).* Less ordered structure of amorphous and glassy materials in comparison to the crystalline state results directly in much stronger distortion of the lattice waves in the form of inelastic (structure) scattering (which is connected with the exchange of energy). Consequently, this leads to a distinct decrease in thermal conductivity and its different variability with temperature. This may be seen from experimental data for amorphous and crystalline sulphur, Fig. 2, and from Berman's data for quartz crystal and quartz glass.

According to Klemens (1951), if the inelastic structure scattering is the only process limiting heat conduction, thermal conductivity of such material should be proportional to T , or

$$k_{(\text{glass})} = ET \dots \dots \dots [24]$$

where E is a numerical constant. Substituting known conductivities and specific heats of glasses into the simplified Debye formula Equation [2], Kittel (27) found that the mean path l is independent of temperature at room temperature and above; i.e., the thermal conductivity of glasses varies with temperature in the same manner as C_v or in accordance with Equation [24].

It has been observed (Bijl, 1949), however, that in the vicinity of the absolute zero (between 1.4 and 3 K) where the specific heat should be proportional to T^3 , the thermal conductivity of glasses varies as $T^{1/2}$. At slightly higher temperatures, say, about 10 to 20 K, the thermal conductivity of glasses varies little with temperatures throughout several degrees and then increases according to Equation [24]. The explanation of this phenomenon is given by Klemens and others. At high frequencies of atomic vibrations characteristic for high temperatures, the wave lengths of the waves carrying the thermal energy are relatively short as compared to the lattice dimensions. Likewise, temperature has little effect on scattering of the waves. For this reason the mean free path l , Equation [2], is independent of frequencies and consequently of temperature. At lower temperatures, however, these frequencies distinctly decrease and the wave lengths increase. When the wave length is large as compared to the lattice dimen-

sions, the disorder of the glass structure has smaller scattering effect and the mean free path increases. A general thermal conductivity-temperature curve for glasses is shown in Fig. 5. Its shape is in accordance with a theoretical treatment by Klemens and with the experimental results of Berman (1951) on quartz glass.

Other experimental data on the thermal conductivity of amorphous and glassy materials available for sulphur, clay bricks, glasses, etc., measured in certain ranges of temperature only, are in full agreement with presented theory. Their thermal conductivity increases with temperature as it is predicted either by inelastic structure scattering of the lattice waves or directly by the dynamic theory of heat.

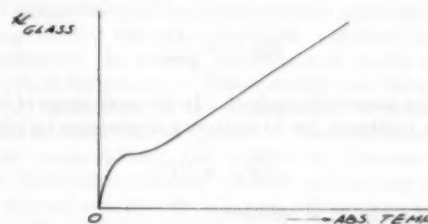


FIG. 5 THERMAL CONDUCTIVITY OF GLASS VERSUS TEMPERATURE

(c) *Refractory Materials.* Variation of the thermal conductivity of refractory materials with temperature depends upon their composition. Those containing more crystalline constituents will show decrease of k with temperature, whereas for those with predominant glassy substances the thermal conductivity will increase with temperature. This fact explains why the thermal conductivity of kaolin firebricks increases with temperature and that of magnesite and silicon-carbide bricks drops with temperature. It also explains different variabilities of k with temperature in a refractory material of the same chemical composition. The ratio of crystalline to glassy constituents is a decisive factor.

(d) *Mixtures of Amorphous Materials.* If a mixture of amorphous materials forms a heterogeneous system, its thermal conductivity may be considered in first approximation as an additive property and calculated by summation of the contributions of constituents.

Solids. Since the thermal conductivity of individual amorphous materials increases with temperature, their mixture should reveal the same behavior. This has been proved experimentally for various kinds of clay bricks (13).

Solid-Liquid Systems. For heterogeneous mixtures of water and amorphous solids, the thermal conductivity increases with temperature (k of water increases up to 248°F). This has been found experimentally for various kinds of wood where "the increase of k with temperature amounts to about 0.1 per cent per degree Fahrenheit" (24).

Since for the majority of liquids k decreases with temperature, it is possible that a heterogeneous mixture of such a liquid and amorphous solid with small temperature coefficient of k will show k of the mixture nearly independent of or even decreasing with temperature.

Solid-Gas Systems. To this class belong nearly all building and insulating materials. Since k of both constituents increases with temperature, the thermal conductivity of solid-gas systems always increases with temperature (38).

In solid-air systems, the thermal conductivity at constant temperature is a function of apparent density (38). This proves that the thermal conductivity of heterogeneous systems is an additive property.

Metals and Semiconductors of Electricity

(a) *Pure Metals.* Since, according to Equation [4], heat in pure metals is conducted independently by the lattice waves and electrons, the dependence of various scattering processes on temperature for both heat carriers must be taken into account. The lattice conductivity as a function of temperature for crystalline solids was discussed in the preceding paragraph (Equations [21] and [23]) and is represented graphically in Fig. 4.

In pure metals, the additional resistance to heat conduction is caused by scattering of phonons by electrons and electrons by phonons. According to Laredo (29) resistance introduced by scattering of phonons by electrons at low temperatures ($T < 0.1\theta$) is inversely proportional to the squared absolute temperature or

$$R_{ph-el} = \frac{1}{FT^2} \quad [25]$$

where F is a numerical constant. In the same range of temperatures, the resistance due to scattering of electrons by phonons is

$$R_{el-ph} = GT^2 \quad [26]$$

where G is a numerical constant.

Since in metals

$$k_t = k_l + k_s = \frac{1}{R_l} + \frac{1}{R_s} \quad [27]$$

where

$$R_l = R_u + R_B + R_{ph-el} \quad [28]$$

and

$$R_s = R_{el-ph} \quad [29]$$

then at low temperatures the total thermal conductivity of metals is

$$k_t = \frac{1}{R_u + R_B + R_{ph-el}} + \frac{1}{R_{el-ph}} \\ = \frac{1}{AT e^{-\frac{1}{2}\frac{\theta}{T}} + B'LT^3} + \frac{1}{FT^2} + \frac{1}{GT^2} \quad [30]$$

Since in pure metals, good conductors of electricity, the electronic conductivity considerably outweighs the lattice conductivity as a first approximation, we may write

$$k_t = k_s = \frac{1}{GT^2} \quad [31]$$

According to Equation [31], the electronic thermal conductivity of pure metals increases with the drop of temperature and tends toward infinity at the absolute zero. This is a logical consequence of the decrease in the electrical resistance at low temperatures and its drop to zero at the absolute zero. At very low temperatures, however, tiny amounts of impurities, which are always present in investigated samples, scatter electrons according to $1/T$ law (29) or

$$R_{el-imp} = \frac{M}{T} \quad [32]$$

where M is a numerical constant depending upon the amount of impurity. Consequently, for pure metals (not absolutely pure!) the electronic conductivity at low temperatures should be expressed as

$$k_s = \frac{1}{R_s} = \frac{1}{GT^2 + \frac{M}{T}} \quad [33]$$

This equation shows a maximum which for the same metals is shifted to lower temperatures as the purity is increased. From the experimental results the coefficients G and M could be evaluated (33).

In poor conductors of electricity, however, the lattice conductivity cannot be neglected and modified Equation [30] must be used

$$k_t = k_l + k_s = \frac{1}{\frac{1}{AT} e^{-\frac{1}{2}\frac{\theta}{T}} + \frac{1}{B'LT^3} + \frac{1}{FT^2}} + \frac{1}{GT^2 + \frac{M}{T}} \quad [30a]$$

Obviously, the resistance offered to heat conduction by scattering of phonons by chemical impurities, Equation [18], should be included in the lattice-conductivity term in Equation [30a]. It is neglected here because at low temperatures the lengths of heat-carrying waves (about 10^{-4} cm) are greater than the impurity scattering centers and the resulting resistance is insignificant.

The thermal conductivity of metals which are poor conductors of electricity is quite a complex function of temperature, but its mathematical analysis is not too difficult. Such analysis indicates an occurrence of a maximum of k_t the position of which is determined by the numerical constants involved in this equation and the value of temperature θ .

At a certain temperature after the maximum, however, the electrical resistance σ of a pure metal becomes proportional to T and the relationship between the thermal conductivity due to electrons and the electrical conductivity may be expressed by the Wiedemann-Franz Law ($k_s = \sigma LoT$ (Equation [5a])). If Lo , the Lorenz number, is assumed constant, then the electronic conductivity in this range of temperatures is independent of temperature. On the other hand, in the same range of temperatures the lattice conductivity becomes proportional to $1/T$, Equation [23]. Consequently

$$k_t = k_l + k_s = \frac{1}{A'T} + \text{const.} \quad [34]$$

or the total thermal conductivity of pure metals will decrease slightly with temperature due to the contribution of the lattice conductivity.

Experimental data on thermal conductivity of metals at low temperatures are available for 34 metallic elements (32). All of them show the characteristic maximum (in most cases between 5 and 30 K). The position of the maximum with respect to temperature and the numerical value of k at the maximum are very sensitive for impurities (constant M in Equation [32]) and may vary for different samples of the same metal.

(b) *Alloys.* Chemical impurities have a pronounced influence on the thermal conductivity of metals and its variability with temperature. They scatter the electrons according to $1/T$ law, Equation [32], and the phonons according to T -law, Equation [18]. The last scattering in low temperatures is negligible. The influence of temperature on the thermal conductivity of alloys depends upon their structure.

Solid-Solution Alloys. The effect of composition on the thermal conductivity of solid-solution alloys was already discussed from the point of view of scattering of phonons and electrons by the presence of foreign atoms in the lattice, Fig. 1.

Since in alloys, as in pure metals, heat is conducted by the lattice waves and electrons, a basic mathematical expression for the thermal conductivity of solid-solution alloys in low temperatures

will be the same as for pure metals which are poor conductors of electricity, Equation [30a]. If the amount of impurity is very small, the k_i of solid-solution alloys shows a maximum as in the case of pure metals. Increasing impurity (manifested by an increase of the numerical constant M) lowers the k_i - T curve and shifts the maximum toward higher temperatures. This can be seen from experimental data of Estermann and Zimmerman (18) on pure and impure germanium in low temperatures. "Impure" germanium contained 0.006 atomic per cent of aluminum.

A large amount of impurity considerably raises the thermal resistance due to scattering of electrons by impurity which becomes a predominant factor in Equation [30a]. Consequently it simplifies to

$$k_i = \frac{T}{M} \dots \dots \dots [35]$$

Equation [35] shows that the thermal conductivity of solid-solution alloys at low temperatures increases proportionately with the absolute temperature and a maximum does not occur.

The variation of k with temperature of solid-solution alloys at higher temperatures also depends upon the amount of impurity (composition). If M is very small and a maximum in low temperatures is preserved, then k will fall with temperature toward the melting point. If M is large, the structure of solid-solution alloys becomes very disordered and scattering of phonons resembles that in glassy materials, Equation [24]. In other words, the lattice conductivity increases with temperature. Assuming the validity of the Wiedemann-Franz law in this range of temperature, we may conclude as a first approximation that the electronic conductivity remains constant. Consequently

$$k_i = E'T + \text{const.} \dots \dots \dots [36]$$

where E' is a numerical constant. Equation [36] shows that k_i of solid-solution alloys increases with temperature also at high temperatures.

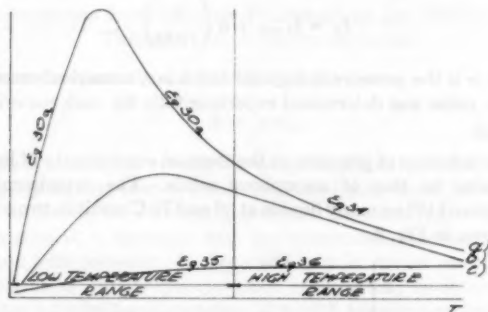


FIG. 6 THERMAL CONDUCTIVITY-TEMPERATURE CURVES FOR PURE AND IMPURE METALS AND SOLID-SOLUTION ALLOYS

In Fig. 6 are sketched three thermal conductivity-temperature curves for low and high temperatures. Curve *a* for a pure metal (containing a very small amount of impurities — M very small) shows a maximum in low temperatures, then falls to melting point in accordance with Equation [34]. In curve *b* for impure metals (M small) the maximum still occurs but is shifted to higher temperatures. Curve *c* is typical for solid-solution alloys.

Eutectic Alloys. In binary eutectic alloys heat is conducted independently by both types of crystals and the total thermal conductivity is the sum of both contributions. Since the eutectic alloys are only mechanical mixtures of different types of crystals, variation of their thermal conductivity with temperature will be of the same character as that of pure metals, Equations [30a] and [34].

(*c*) **Semiconductors.** Although in semiconductors, heat is conducted by the lattice waves and free electrons, the electronic conductivity is only a fraction of the total conductivity. For this reason the variation of the thermal conductivity in semiconductors with temperature will follow the pattern for dielectric solids. This can be seen from experimental results of Powell (40) on various types of graphite.

The number of free electrons in semiconductors rapidly increases with temperature and at very high temperatures the electronic contribution to k_i may outweigh the lattice contribution and the semiconductor will act as a metal. This, however, will not change the shape of k_i - T curve.

Liquids

Lack of proper theory of the liquid structure makes difficult the explanation of the thermal-conductivity variation of liquids with temperature. In general, the volume of liquids increases with the rise of temperature. This indicates that the degree of disorder or randomness in the arrangement of molecules in liquids also increases with temperature. Since according to the concept of order-disorder the higher the disorder among molecules, the lower the thermal conductivity, one may conclude that the thermal conductivity of liquids will decrease with temperature. The available experimental data for the majority of nonmetallic liquids and molten metals at atmospheric pressure are in accordance with this theory because their k decreases with temperature. On the other hand, there are some liquids which show a positive temperature coefficient of k .

Nonmetallic Liquids. (*a*) **At 1 Atm Pressure.** Hydrocarbons and nearly all organic liquids under atmospheric pressure show a decrease of thermal conductivity with temperature, but water, aqueous solutions, molten sulphur, and a few organic liquids (glycerin, ethylene glycol) reveal a positive temperature coefficient of k . The explanation of this phenomenon has been sought by many authors. So far, the most convincing argument was presented by Bridgman with his formula, Equation [10], according to which the thermal conductivity of liquids is proportional to the velocity of sound and inversely proportional to the specific volume to 2/3 power. Under atmospheric pressure, the specific volume of all liquids increases with temperature. The velocity of sound, however, decreases with temperature only for organic liquids. For water, it increases with temperature (24). In organic liquids, therefore, both factors act toward the decrease of k with temperature and that is in full accordance with experimental results. In water, however, the rate of increase of the velocity of sound with temperature is higher than that of specific volume to 2/3 power; consequently, the thermal conductivity of water increases with temperature.

(*b*) **Fluids in Saturated State.** Experimental data available for liquids in saturated state, i.e., for water, ammonia, carbon dioxide, sulphur dioxide, and some organic liquids (17), show a general tendency to fall with temperature toward the critical point. This is achieved either by passing through a maximum in liquids in which k increases with temperature at pressure close to atmospheric (water, NH_3 , CO_2) or by a continuous drop in k after melting (methyl chloride). The thermal conductivity of water and water vapor in saturated state is shown in Fig. 7 (17). At the critical point, the thermal conductivity of fluids has a finite value, 0.121 Btu/[hr sq ft (deg F/ft)] for water and 0.034 for carbon dioxide (24). Below the critical point, the thermal conductivity of saturated vapor falls with temperature.

Liquid Metals. In liquid metals heat is conducted by atomic vibrations and free electrons. For this reason the liquid metals show a much higher thermal conductivity than nonmetallic liquids. According to Powell (41), the majority of pure metals in molten state show a decrease of k with temperature. On the

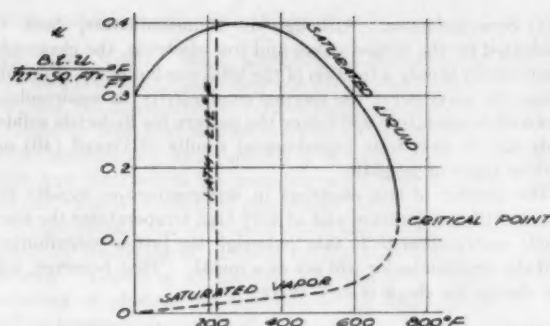


FIG. 7 THERMAL CONDUCTIVITY OF WATER AND WATER VAPOR IN SATURATED STATE

other hand, liquid mercury and cadmium, and four eutectic alloys (Pb-Bi, Sn-Pb, Sn-Zn, and Pb-Sb) have a positive temperature coefficient of k . There is no available theory explaining the variation of k with temperature in liquid metals.

Gases

According to Equation [14], the thermal conductivity of gases is a product of three terms, magnifying factor a , viscosity μ , and specific heat at constant volume C_v . The variation of the thermal conductivity of gases with temperature, therefore, will be a resultant of the changes with temperature of all three terms. Let us consider these changes at atmospheric pressure (a simultaneous influence of temperature and pressure on the thermal conductivity of gases is discussed in the next paragraph).

According to Eucken's formula, factor a is directly proportional to the ratio of specific heats [$a = 0.25(9\gamma - 5)$] where $\gamma = C_p/C_v$ and will decrease, therefore, with temperature. On the other hand, available data on γ for various gases under atmospheric pressure, especially in higher temperatures, show that this decrease is rather small (i.e., about 0.04 per cent per 1 deg F for hydrogen) and may be neglected.

Viscosity of gases increases with temperature and according to W. Sutherland (1893) it may be represented by the following function of absolute temperature

$$\mu = \frac{\mu_0 \sqrt{T}}{1 + \frac{T}{C}} \quad [37]$$

where μ_0 and C are constants for the gas, and T is the absolute temperature.

Specific heat at constant volume increases with temperature in various degrees at various temperatures (20), but for short temperature intervals it may be assumed as a linear function of temperature, or

$$C_v = C_{v0}(1 + NT) \quad [38]$$

where N/C_{v0} is the temperature coefficient of C_v .

Substitution of Equations [37] and [38] into Equation [14] results in a formula for the thermal conductivity of gases as a function of temperature

$$k = a\mu C_v = a \times \frac{\mu_0 \sqrt{T}}{1 + \frac{T}{C}} C_{v0}(1 + NT)$$

$$= a\mu_0 C_{v0} \frac{\sqrt{T}(1 + NT)}{1 + \frac{T}{C}} = k_0 \frac{(1 + NT)T^{3/2}}{T + C} \quad [39]$$

According to Jakob (24) this equation with proper constants for a given gas "represents the known experimental data with an accuracy of 0 to 6 per cent in the range of -180°C to $+200^\circ\text{C}$."

In high temperatures, however, this type of formula is not valid. For instance, the experimental results of Stops (47) on air, nitrogen, and carbon dioxide over the temperature range 0 to 1000°C inclined toward a cubic equation of temperature. The Sutherland type of equation could be used for air and nitrogen only up to 600°C .

VARIABILITY OF THERMAL CONDUCTIVITY WITH PRESSURE

Solids

Metals. From the theory of the lattice and electronic conductivities in metals it could be expected that the pressure will distort the regularity of the lattice which in turn will increase scattering of phonons and electrons; i.e., will decrease the thermal conductivity. This conclusion is, however, only partially confirmed by experimental observations. The experiments of Bridgman on 11 metals with pressures up to $12,000 \text{ kg/cm}^2$ (10) showed that the thermal conductivity of iron, copper, silver, nickel, platinum, bismuth, and antimony decreases with pressure linearly (at various rates), whereas that of lead, tin, cadmium, and zinc increases in a similar manner. On the other hand, the pressure coefficient of the thermal conductivity of metals is quite small. This last fact shows that pressures used by Bridgman cause only elastic deformation of the metal lattice and as such may hinder or promote the propagation of the lattice waves and the flow of electrons depending upon the structure of the lattice.

Amorphous Solids. Pressure exerted on amorphous solids increases the contact area between molecules and atoms and should, therefore, promote heat conduction by atomic vibrations. This conclusion was confirmed by experiments of Bridgman (10) who has shown that for nonmetallic, amorphous solids (pyrex glass, limestone, talc, etc.) the thermal conductivity increases proportionally to the pressure, according to the formula

$$k_p = k_{1 \text{ atm}} + a \left(\frac{p}{1000} \right) \quad [40]$$

where p is the pressure in kg/cm^2 and a is a numerical constant, whose value was determined experimentally for each material.

Liquids

The influence of pressure on the thermal conductivity of liquids is similar to that of amorphous solids. The experiments of Bridgman (10) on many liquids at 30 and 70°C result in two curves as shown in Fig. 8.

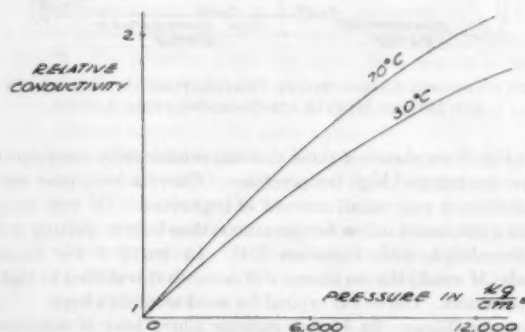


FIG. 8 INFLUENCE OF PRESSURE ON THERMAL CONDUCTIVITY OF LIQUIDS (BRIDGMAN)

Gases

The kinetic theory of gases leads to a general expression for the thermal conductivity of gases

$$k = \frac{1}{3} \bar{u} C_v \dots \dots \dots [11]$$

It indicates that k of the perfect gas is independent of pressure for the following reasons: At constant temperature, neither \bar{u} , the mean velocity of the molecules, nor C_v , the specific heat at constant volume, is affected by pressure. Only l , the average path length between the collisions of molecules, will decrease with pressure. But the pressure will increase the density of the gas or the number of molecules per unit of volume. Since in the perfect gas the decrease in l will be exactly balanced by the increase in number of molecules taking part in the heat transportation, k remains unchanged. However, even for the perfect gas this conclusion is not valid at reduced pressures. Heat conductivity in gases is connected with a large number of molecules at unit volume. At low pressures there are fewer molecules available for the transport of heat, their mean paths are much longer, and collisions less frequent. In this region of pressures the thermal conductivity of the perfect gas is proportional to the pressure.

In real gases, however, there are two additional factors which result in a general dependence of k on pressure. First, C_v of real gases is not constant but increases with pressure. This is quite apparent for polyatomic gases. Also, the molecules of real gases are not hard spherical bodies without volume but possess a finite volume and may have various shapes. For this reason the mean free path between the collisions of molecules in real gases is not equal to the distance between the centers of molecules but somewhat less; this difference increases with pressure.

The influence of pressure on k of real gases is difficult to observe for small pressure differences, especially in the vicinity of 1 atm. In this region, real gases are still close to the perfect gas and their k is practically independent of pressure. Large pressure differences, however, show positive pressure coefficients of k . These coefficients are quite high for heavy gases and vapors with polyatomic molecules.

SIMULTANEOUS INFLUENCE OF TEMPERATURE AND PRESSURE ON THERMAL CONDUCTIVITY OF GASES

Since the thermal conductivity of gases

$$k = a \mu C_v \dots \dots \dots [14]$$

then a simultaneous influence of temperature and pressure on k of actual gases may best be obtained by analyzing the effect of these two factors on the three components of Equation [14].

In general, γ decreases with the temperature increase and increases with pressure. Factor a will vary in similar manner and its variation with temperature and pressure may be obtained by Eucken's formula and variation of γ with temperature and pressure. The data are available for air (16), water vapor, and some other gases.

The viscosity and specific heat at constant volume of gases vary with temperature and pressure in the same manner according to the well-known pattern shown in Fig. 9.

Since the variations of a with temperature and pressure are relatively small, the general shape of curves in Fig. 9 must be preserved for the variation of the thermal conductivity of gases with temperature and pressure. This conclusion was proved experimentally for water vapor by Timrot and Vargaftig (50) whose results, represented graphically, fall into the general pattern of Fig. 9 (24).

At very low pressures, the thermal conductivity of actual gases varies with pressure for the same reason as explained for the perfect gas. According to Jakob (24), k for heavy gases depends upon pressure below 1 mm Hg whereas that for hydrogen and helium, below 20 mm Hg.

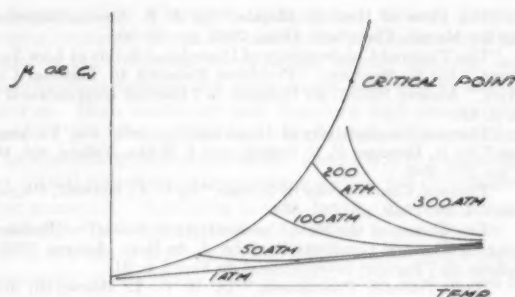


FIG. 9 GENERAL CURVES FOR VARIATION OF μ AND C_v WITH TEMPERATURE AND PRESSURE

SUMMARY

The influence of material structure, temperature, and pressure on the thermal conductivity in the three physical states of matter was discussed and theoretical conclusions were compared with experimental results.

1 Thermal conductivity of matter depends upon its structure to a considerable degree. Structurally, more organized materials show higher thermal conductivity. For this reason it increases in the following succession: Gases—liquids—amorphous and glassy materials—crystals. Solid and molten metals show higher thermal conductivity than dielectric solids and nonmetallic liquids, respectively, because in metals heat conduction by atomic vibrations (the lattice conductivity) is strongly supported by a drift of free electrons.

2 Temperature. For a material appearing in crystalline and amorphous-solid phase, the dependence of k upon temperature at atmospheric pressure throughout the three physical states may be represented by a general pattern, as shown in Fig. 10.

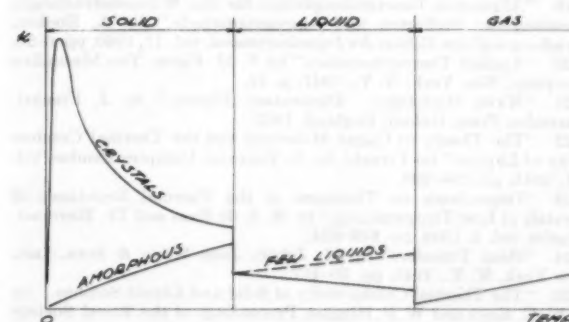


FIG. 10 THERMAL CONDUCTIVITY AS FUNCTION OF TEMPERATURE THROUGHOUT THREE PHYSICAL STATES

The general trend in variation of k with temperature has been proved by theory and experiment. From this graph and known structure of solid sample, the variation of its thermal conductivity with temperature may be predicted (refractory materials, eutectic alloys, amorphous heterogeneous mixtures, etc.).

3 Pressure. Generally speaking, k of all materials increases with pressure. The only deviation is shown by some metals.

BIBLIOGRAPHY

- 1 "An Analysis of Heat Transfer Through Thermal Insulating Materials," by E. A. Allcut, ASME Proceedings of the General Discussion on Heat Transfer, September 11-13, 1951, pp. 232-235.
- 2 "Thermal Conductivities of Pure Metals at Low Temperatures. I. Aluminum," by F. A. Andrews, R. T. Webber, and D. H. Spohr, *Physical Review*, vol. 84, 1951, pp. 994-996.

- 3 "The Flow of Heat in Metals," by J. B. Austin, American Society for Metals, Cleveland, Ohio, 1942, pp. 35-50.
- 4 "The Thermal Conductivity of Disordered Solids at Low Temperatures," by R. Berman; "Problems Relating to Thermal Conductivity," Annexe 1952-1, au Bulletin de l'Institut International du Froid, p. 13.
- 5 "Thermal Conductivity of Dielectric Crystals: the 'Umklapp' Process," by R. Berman, F. E. Simon, and J. Wilks, *Nature*, vol. 168, 1951, pp. 277-280.
- 6 "Thermal Conductivity of Metals," by C. C. Bidwell, *Physical Review*, vol. 58, 1940, pp. 561-564.
- 7 "The Theory of the Heat Conductivity in Solids"—"Problems Relating to Thermal Conductivity," by J. de Boer, Annexe 1952-1, au Bulletin de l'Institut International du Froid, p. 21.
- 8 "Heat Transfer Phenomena," by R. C. L. Bosworth, John Wiley & Sons, Inc., New York, N. Y., 1952.
- 9 "The Thermal Conductivity of Liquids Under Pressure," by P. W. Bridgman, Proceedings of the American Academy of Arts and Sciences, vol. 59, 1923, pp. 141-169.
- 10 "The Physics of High Pressure," by P. W. Bridgman, G. Bell and Sons, London, England, 1949, pp. 307-326.
- 11 "Conductibilité Calorifique et Viscosité des Liquides Monoatomiques," by L. Brillouin, *Comptes Rendus*, vol. 159, 1914, pp. 27-30.
- 12 "Note on the Conduction of Heat in Crystals," by H. B. G. Casimir, *Physica*, vol. 5, 1938, pp. 495-500.
- 13 "Manual for Process Engineering Calculations," by L. Clarke, McGraw-Hill Book Company, Inc., New York, N. Y., 1947, pp. 234-235.
- 14 "Vortrage über der Kinetische Theorie der Materie und der Elektrizität," by P. Debye, Teubner, Berlin, Germany, 1941.
- 15 "Note on a Method of Estimating the Prandtl Number of Liquids," by K. G. Denbigh, *Journal of the Society of Chemical Industry*, vol. 65, 1946, pp. 61-63.
- 16 "Chemical Engineering Thermodynamics," by B. F. Dodge, McGraw-Hill Book Company, Inc., New York, N. Y., 1944, p. 270.
- 17 "Introduction to the Transfer of Heat and Mass," by E. R. G. Eckert, McGraw-Hill Book Company, Inc., New York, N. Y., 1950, pp. 265-272.
- 18 "Effect of Lattice Defects on the Thermal Conductivity of Certain Alloys at Low Temperatures," by I. Estermann and J. E. Zimmerman—"Problems Relating to Thermal Conductivity," Annexe 1952-1 au Bulletin de l'Institut International du Froid, p. 33.
- 19 "Allgemeine Gesetzmässigkeiten für das Wärmeleitvermögen Verschiedener Stoffarten und Aggregatzustände," by A. Eucken, *Forschung auf dem Gebiete des Ingenieurwesens*, vol. 11, 1940, pp. 6-20.
- 20 "Applied Thermodynamics," by V. M. Faires, The Macmillan Company, New York, N. Y., 1947, p. 32.
- 21 "Wave Mechanics: Elementary Theory," by J. Frenkel, Clarendon Press, Oxford, England, 1932.
- 22 "The Theory of Caged Molecules and the Thermal Conductivity of Liquids" (in French), by A. Fournier, *Comptes Rendus*, vol. 221, 1945, pp. 294-296.
- 23 "Dependence on Thickness of the Thermal Resistance of Crystals at Low Temperatures," by W. J. de Haas and Th. Biermass, *Physica*, vol. 5, 1938, pp. 619-624.
- 24 "Heat Transfer," by M. Jakob, John Wiley & Sons, Inc., New York, N. Y., 1949, pp. 70-117.
- 25 "The Thermal Conductivity of Solid and Liquid Sulphur," by G. W. C. Kaye and W. F. Higgins, Proceedings of the Royal Society of London, series A, vol. 122, 1929, pp. 633-646.
- 26 "Free Volumes and Free Angle Ratios of Molecules in Liquids," by J. F. Kincaid and H. Eyring, *Journal of Chemical Physics*, vol. 6, 1938, pp. 620-629.
- 27 "Interpretation of the Thermal Conductivity of Gases," by C. Kittel, *Physical Review*, vol. 75, 1949, pp. 972-974.
- 28 "The Thermal Conductivity of Dielectric Solids at Low Temperatures," by P. G. Klemens, Proceedings of the Royal Society of London, series A, vol. 208, 1951, pp. 108-133.
- 29 "Recent Work on Thermal Conductivity of Metals at Very Low Temperatures," by S. J. Laredo—"Problems Relating to Thermal Conductivity," Annexe 1952-1 au Bulletin de l'Institut International du Froid, p. 61.
- 30 "The Thermal Conductivity of Metals," by R. E. B. Makinson, Proceedings of the Cambridge Philosophical Society, Cambridge, England, vol. 34, 1938, pp. 474-497.
- 31 "Heat Transmission," by W. H. McAdams, McGraw-Hill Book Company, Inc., New York, N. Y., third edition, 1954, p. 28.
- 32 "Heat Conductivity of Metals at Low Temperatures," by K. Mendelssohn—"Problems Relating to Thermal Conductivity," Annexe 1952-1 au Bulletin de l'Institut International du Froid, p. 67.
- 33 "The Thermal Conductivity of Metals at Low Temperatures. I. The Elements of Groups 1, 2, and 3," by K. Mendelssohn and H. M. Rosenberg, Proceedings of the Physical Society, series A, vol. 65, 1952, pp. 385-388.
- 34 "Thermal Conductivity of Liquids," by G. Palmer, *Industrial and Engineering Chemistry*, vol. 40, 1948, pp. 89-92.
- 35 Paschki, *Journal of the Russian Chemical Society*, 1915, p. 276.
- 36 "Zur Kinetischen Theorie der Wärmeleitung in Kristallen," by R. Peierls, *Annalen der Physik*, vol. 3, 1929, pp. 1055-1101; and "Quelques Propriétés Typiques des Corps Solides," *Annales de l'Institut Henri Poincaré*, vol. 5, 1935, pp. 177-222.
- 37 "Zur Kinetischen Theorie der Wärmeleitung in Kristallen," by R. Peierls, *Annalen der Physik*, vol. 3, 1929, pp. 1055-1101.
- 38 "Chemical Engineers' Handbook," by J. H. Perry, McGraw-Hill Book Company, Inc., New York, N. Y., third edition, 1950, p. 458.
- 39 "Thermal Conductivity of Dielectrics at Temperatures Lower Than That of Debye," by J. Pomeranchuk, *Journal of Physics, USSR*, vol. 6, 1942, pp. 237-250.
- 40 "Thermal Conductivities of Metallic Conductors, and Their Estimation," by R. W. Powell, ASME Proceedings of the General Discussion on Heat Transfer, September 11-13, 1951, pp. 290-295.
- 41 "Thermal Conductivities of Molten Metals and Alloys," by R. W. Powell, *Journal of the Iron and Steel Institute*, London, England, vol. 162, 1949, pp. 315-324.
- 42 "The Variation With Temperature of the Thermal Conductivity . . . of Some Micas," by R. W. Powell and E. Griffiths, Proceedings of the Royal Society of London, England, series A, vol. 163, 1937, pp. 189-198.
- 43 "Diffusion, Thermal Conductivity, and Viscous Flow of Liquids," by R. E. Powell, W. E. Roseveare, and H. Eyring, *Industrial and Engineering Chemistry*, vol. 33, 1941, pp. 430-435.
- 44 "Thermal Conductivity of Liquids," by M. Rama Rao, *Indian Journal of Physics*, vol. 16, 1942, pp. 161-167.
- 45 "Thermal Conductivity of Liquid Metals," by M. Rama Rao, *Indian Journal of Physics*, vol. 16, 1942, pp. 155-159.
- 46 "New Thermal Conductivity Measurements for Organic Liquids," by L. Riedel, *Chemie Eng. Tech.*, vol. 23, 1951, pp. 321-324.
- 47 "Effect of Temperature Upon the Thermal Conductivity of Gases," by D. W. Stops, *Nature*, vol. 164, 1949, pp. 966-967.
- 48 "Thermal Conductivity of Liquids," by J. F. D. Smith, *Industrial and Engineering Chemistry*, vol. 22, 1930, pp. 1246-1251.
- 49 "Dimensional Analysis Applied to the Thermal Conductivity of Liquids," by J. F. D. Smith, *Industrial and Engineering Chemistry*, vol. 23, 1931, pp. 416-419.
- 50 "Thermal Conductivity of Liquids," by J. F. D. Smith, *Trans. ASME*, vol. 58, 1936, pp. 719-725.
- 51 "The Theory of the Transport Phenomena in Metals," by E. H. Sondheimer, Proceedings of the Royal Society of London, England, series A, vol. 203, 1950, pp. 75-98.
- 52 "The Thermal Conductivity and Viscosity of Steam at High Temperature and Pressures," by D. L. Timrot and N. B. Vargaftig, *Journal of Physics, USSR*, vol. 2, 1940, p. 101.
- 53 H. F. Weber, *Wiedemann Annalen*, vol. 48, 1893, p. 173.
- 54 "Heat," by A. G. Worthing and D. Halliday, John Wiley & Sons, Inc., New York, N. Y., 1948, p. 102.

Discussion

VICTOR PASCHKIS.^a Information on thermal conductivity is always welcome. In this paper there are several comments that come to mind and on which additional information from the author would be welcome.

1 In the section on glassy materials the author quotes values of "thermal conductivity of quartz glass." Heat transfer through glass is always by conduction and radiation. Determination of conductivity by conventional means yields an apparent conductivity rather than the direct conductivity. Does the author refer to apparent or direct values and, if the latter, how have they been determined?

2 With reference to the schematic Fig. 10, it may be of interest that for carbon steel the conductivity of the liquid near the melting point seems to be one half of that of the solid near the melting point. Dr. J. B. Austin suggested this relationship some

^a Technical Director, Heat and Mass Flow Analyzer Laboratory, Columbia University, New York, N. Y. Mem. ASME.

10 years ago in a personal communication to the writer. It has been used repeatedly in computations of solidification problems by means of electric analogy. Such computations have subsequently been compared with temperature measurements and have been found correct. A similar ratio has been found for aluminum by the same method: computing the solidification pattern with this assumption, and comparing the result with temperature measurements.

3 In the last section of the paper the author speaks about the variation of conductivity with pressure. It is known that the conductivity of gases increases at high pressures, but little quantitative information seems available. In view of the increase in use of high-pressure apparatus and of the fact that the conductivity of insulating materials is controlled by the conductivity of the gas in the pores of the material, accurate knowledge of the pressure dependence of gases at high temperature would be of interest.

AUTHOR'S CLOSURE

The data on thermal conductivities of quartz crystal and quartz glass are taken from the paper of Berman (4) in which the methods used for evaluation of thermal conductivities are not disclosed. Probably they could be obtained directly from Dr.

R. Berman (The Clarendon Laboratory, Oxford, England). However, it seems that in the case of the particular experiments mentioned the correction for radiation, usually required, would be unimportant because the experiments were run at low temperatures. High emissivity and, therefore, high absorptivity of quartz (31, p. 478) is another factor diminishing this correction.

The interesting observation of Mr. Paschkis is in general agreement with the available data for some metals, but exceptions are rather numerous. According to J. B. Austin (3), "at the melting temperature the conductivity of a liquid metal is from one third to two thirds that of a solid." Bismuth for which thermal conductivity in the liquid state is higher than that in the solid state (see Fig. 3) is considered to be an exception.

We heartily agree with Mr. Paschkis that additional data on thermal conductivity of gases at high pressure would be very helpful. The experiments of Timrot and Vargaftig (50) on water vapor follow a general pattern deducible from Equation [14]. Since the preparation of this paper, some additional information^a on the subject became available.

^a "Thermal Conductivity of Gases," by F. G. Keyes, Trans. ASME, vol. 76, 1954, pp. 809-816.

Also, "Thermal Conductivity of Gases," by F. G. Keyes, ASME Paper No. 54-A-235 (to be published).

The Ultrasonic Measurement of Hydraulic Turbine Discharge

By R. C. SWENGEL,¹ W. B. HESS,² AND S. K. WALDORF³

To obtain the basic discharge data necessary for efficient hydroelectric operation, a new ultrasonic flowmeter has been developed and applied to the measurement of flow in large water passages. During the development period, extensive tests under varied conditions of flow were made in a passage, 5 in. \times 9 in., using a weighing tank as the absolute standard. Accuracies of 1 per cent or less were achieved regularly with appreciably turbulent and non-uniform flow distributions, using river water of varying turbidity. Installation in a turbine intake, 16 ft \times 25 ft, has demonstrated the practicality of the method and equipment.

INTRODUCTION

A NUMBER of basically different types of flowmeters or methods for measuring the flow of liquids and gases are in existence today. Some of these methods provide a direct and absolute measurement while others require calibration from which a coefficient is derived and applied in the determination of the quantity of flow. While many of the methods are practical for specific conditions, there are situations in which, for one reason or another, all of them may fail to offer a satisfactory solution. The measurement of large quantities of water is one condition where very few methods can be applied. The limits of accuracy or the cost of instrumentation and installation often preclude the use of those that are applicable.

Low-head hydroelectric plants, particularly those with fixed or adjustable-blade propeller types of turbines, present a problem in flow measurement. Few turbines of this type have been tested for performance characteristics in their installed location. Aside from the cost of applicable methods, field testing has been limited by the short length of water passages and the geometry of the intakes that are characteristic of these installations.

Appreciable economies can be realized from the operation of large turbines when reliable performance data are available. Sometimes too much emphasis is placed upon the results of field tests as the criterion of guaranteed performance rather than as a source of information for most efficient operation.

Performance data on low-head developments, particularly in the United States, are frequently derived from laboratory tests on a homologous model with the data stepped up to prototype by basic dimensional relationships plus empirical or experience coefficients to compensate for other less tangible factors that vary with the size ratio. Actual field tests indicate that data derived in this manner may result in misleading operating procedures. Piezometer combinations in the turbine intake or scroll case have

been used quite extensively as an index for establishing the shape of performance curves and, in the case of adjustable-blade turbines, for obtaining the best efficiency relationships between gate and blade positions. Index tests, however, do not provide accurate quantitative data.

In the case of several units in one plant, an economic gain is realized from preferential operation of the more efficient units. With units of identical design, variations in performance between individual units may seem insignificant; however, field testing has shown differences of 1 to 2 per cent in efficiency which could be the source of appreciable losses if these differences were ignored. With more than one plant on a stream, gains can be achieved when reliable performance data are available as a basis for more economical distribution of loading between plants and better utilization of storage.

The cited advantages accruing from economical operation provided the incentive to search for an improved method of flow measurement. The literature search revealed several good texts (1 to 5)⁴ and considerable previous work in the general field of ultrasonics, but left unsolved serious difficulties which prevented its practical application to flow measurement. One of the authors of this paper conceived a new ultrasonic method which overcomes prior difficulties (6, 7) and which has been under continuous development and testing since 1947.

Ultrasonic Flowmeter. The new ultrasonic method is based upon the measurement of fundamental units of time and length and requires no calibration by other means prior to its use. The measurement of the difference in propagation velocities of ultrasonic compression waves with and against the direction of flow, as in earlier methods (8 to 12), is the basic phenomenon from which the flow velocity is determined. The principal innovation is the use of fixed elements which function both as transmitter and receiver.

In order to measure the average velocity of flow over the entire area of a rectangular section, newly developed antenna-type transducers are employed. These transducers take the form of straight cylindrical rods which transform electrical waves into mechanical waves and vice versa.

Tests. Extensive tests in a steel conduit having a rectangular cross section 5 in. \times 9 in. have proven the method practical and capable of a high degree of accuracy, affected only a negligible amount by turbidity, temperature, and departures from uniform velocity distribution normally encountered in straight flow sections. A few large-scale tests were made in a turbine intake 16 ft \times 25 ft, where the ultrasonic flowmeter yielded very consistent data.

A rather complete analysis and some experiments indicate that the method can be used to measure fluid flow in open channels, round pipes, and passages having other than rectangular cross sections, although the precision with which such measurements can be made is yet to be determined. Early in the development, experiments were conducted with transducers mounted on the outside of the steel conduit. These trials indicated that such an arrangement is workable, utilizing a variation of the wave-guide techniques common to microwave radio communication.

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

¹ Consultant, York, Pa.

² Senior Engineer, Safe Harbor Water Power Corporation, Conestoga, Pa. Mem. ASME.

³ Engineer of Research, Pennsylvania Water & Power Company, Lancaster, Pa.

Contributed by the Hydraulic Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 9, 1954. Paper No. 54-A-54.

This paper is one of a planned series of three on the ultrasonic flowmeter and includes a general description of the method and representative hydraulic test data. The second paper is to be written for the American Institute of Electrical Engineers and will cover the theory of the method and details of the electronic equipment. The development and characteristics of the novel antenna-type transducers is to be the subject of a third paper.

PRINCIPLES OF THE ULTRASONIC METHOD

In a still medium, compression waves having a fixed frequency radiate from a stationary point source as shown in Fig. 1(a). The propagation velocity being the same in all directions, the wave length is constant and obeys the following law

$$\lambda = \frac{V}{f} \quad [1]$$

where

λ = wave length, ft

V = absolute propagation velocity, fps

f = frequency, cycles per sec (cps)

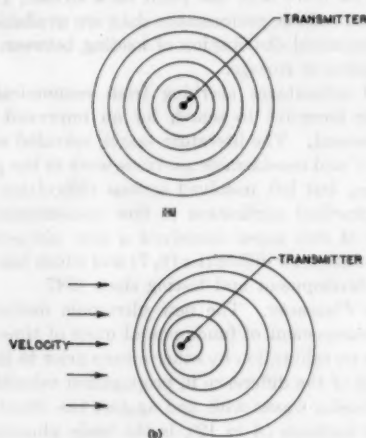


FIG. 1 IDEALIZED WAVE PATTERNS
(a—Stationary medium. b—Medium moving with uniform velocity.)

Fig. 1(b) shows the wave displacement when the medium moves uniformly at constant velocity. The wave length is increased in the direction of flow and decreased against the flow. The basic Equation [1] can be modified to include the effect of the velocity of flow v

$$\lambda = \frac{V \pm v}{f} \quad [2]$$

Since V has a value of about 4800 fps in water, and the velocity of flow v in low-head turbines is often as small as 1 fps, it is necessary to measure changes in sound velocity due to flow of the order of one part in 4800. Further, it is desirable to measure such a change to within 1 per cent or less, which requires that a change in sound velocity of about 1 part in 480,000 be measured. As shown in Equation [2], a change in flow velocity v produces a change in λ proportional to the change in $V \pm v$. To amplify a change in λ , the frequency f is chosen high enough to produce a train of about 100 waves in the space between the transmitter and receiver. The wave at the receiver is advanced or retarded with respect to the transmitter by an amount which is the summation of the changes in length of the individual waves in the train. The magnitude of this wave shift is determined electrically by measuring phase angles between the wave being transmitted and the wave being received. Other known techniques could have

been used, because this determination is essentially a measurement of the change in transit time due to flow velocity.

A given change in either V or v will produce an identical change in λ . This imposes a serious handicap because a 1 deg C change in temperature changes the speed of sound in water about 9 fps, which is enough to mask the flow velocity to be measured. A still more serious handicap has been a small change in the spacing between transducers occurring during measurement, because such a change could be larger than the minute change in wave length which is the basis of the ultrasonic measurement. Earlier workers reduced the effect of temperature by combinations of transmitting and receiving transducers, but failed to overcome the considerable errors introduced by slight changes in transducer spacing. Both of these sources of error, in effect, have been eliminated in the new method by reversing the direction of transmission between a pair of transducers, the maximum benefit being derived when transmissions are made in both directions simultaneously. Since the magnitudes of the errors are related to the rates of change of temperature and transducer spacing and these rates are very small in turbine testing, it has been found practical to use the less complicated alternate upstream and downstream transmission technique.

The following formula is used to calculate flow velocity

$$v = \frac{V^2}{fd} \times \frac{\Delta}{720} \quad [3]$$

where

v = flow velocity, fps

V = absolute propagation velocity, fps

f = frequency of ultrasonic waves, cps

d = projection of transducer spacing on axis of flow, ft

Δ = difference between the phase angles measured when transmitting with and against flow, deg

Although the derivation of Equation [3] includes a mathematical approximation, the accuracy of the formula is appreciably greater than that attainable in the physical and electrical measurements. It will be noted that there are no empirical coefficients.

EQUIPMENT

A block diagram of the equipment is shown in Fig. 2. The antenna-type transducers are mounted on opposite walls of the measuring section and displaced a definite distance along the flow axis. The transducer assembly consists of a barium-titanate

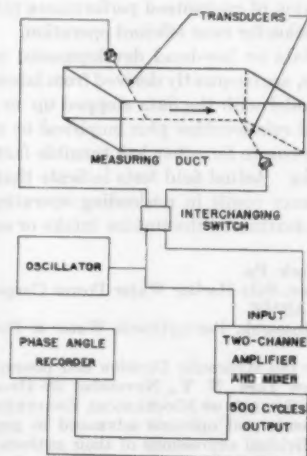


FIG. 2 BLOCK DIAGRAM OF EQUIPMENT USED

transducing element bonded to a solid metal horn, which serves as a mechanical impedance-matching transformer between the transducing element and a slender stainless-steel rod. The barium-titanate element, the transformer, and the connecting electrical cables are moistureproofed by enclosing them in pressurized jackets.

When the transducing element is driven by the alternating electric signal, it vibrates longitudinally at the same frequency as the signal and sends compression waves down the rod, thence into the water. This action takes place in the reverse order when a transducer is used as a receiver.

The antennalike transducers have been designed to overcome three major problems that may be encountered in large turbine intakes:

- 1 Obtaining the true average of velocity over the entire area of a large flow section. The new transducers do this quite successfully by spreading the transmitted ultrasonic energy evenly over the entire area and by vector addition producing a received signal whose vector relationship to the transmitted signal is a measure of the average flow velocity over the entire area.

- 2 The reduction of the effects of echoes or reflections from upstream and downstream surfaces. Since reflected signals may pass through flow having velocities differing from those in the measuring section, an error in phase angle can occur. These effects are attenuated to a negligible level by terminating the transducer antenna rods, thereby producing an effective beam width so narrow that, as installed, only by the direct path between transducers is the signal transmitted or received. To align the ultrasonic beam with the direct path, the transducers are tilted at an angle dependent upon the width of the flow section, the transducer spacing, and the speeds of sound in the transducer rods and in water. Terminated rods were used only for the tests in the turbine intake.

- 3 The problem of interference with normal flow was eliminated by using $\frac{1}{8}$ -in.-diam rods recessed in the walls.

The mechanical interchanging switch is motor driven and interchanges the connections to the transducers every 5 sec, thereby reversing their functions at a predetermined cyclic rate of six times per min.

The oscillator supplies alternating current to the transmitting transducer and a sample of this current to the measuring circuit as an in-phase reference voltage. The frequency of the oscillator is chosen to produce adequate phase-angle deflections for the velocities to be measured, the transducer spacing, and the sensitivity and precision of the measuring circuit.

The two-channel amplifier and mixer amplify the reference signal from the oscillator and the signal from the receiving transducer, electronically reduce their frequency to the audible range where the phase-angle recorder can measure their phase relationship precisely.

The phase-angle recorder consists of two principal parts:

- 1 The phase-angle-measuring circuit which translates the phase angle between the reference and receiving transducer signals into a proportional direct current.

- 2 The Minneapolis-Honeywell high-speed electronic recorder which automatically records the direct current as phase angle in degrees. The record is made on a strip chart 10 in. wide.

PERFORMANCE TESTS IN 5-IN. \times 9-IN. FLOW SECTION

Test Conditions. The flow section was rectangular, 5 in. \times 9 in., and 3 ft long. Full-section flow was maintained at all velocities by locating the regulating valve beyond the discharge end of the test section. A weighing tank of 4100 gal capacity, with its associated beam scale, was used as the absolute standard of accuracy in the initial tests which established the validity of the

method. To facilitate additional testing of the flowmeter under many different conditions, a venturi meter was utilized as the reference standard after being calibrated very carefully with the weighing tank. The oscillator frequency was 500 kc.

A standard procedure was followed for each comparison test:

- 1 Flow was adjusted to a predetermined rate.
- 2 Water temperature was measured to the nearest 0.1 deg C.
- 3 Twenty pairs of ultrasonic readings (upstream and downstream) were recorded automatically.
- 4 The amount of water passing through the section during the period of time the ultrasonic readings were recorded was collected in the tank and the volume determined by weighing. When the venturi meter was used, visual readings from the manometer were recorded for each pair of ultrasonic readings.
- 5 Water temperature again was read.

A complete series of comparison tests was made over the entire range of flow velocities before making any computations. This was done to obtain purely objective data. Before and after each series, the reliability of the electronic instrumentation was checked.

Turbidity Effects. Untreated Susquehanna River water was used for all tests. From time to time, its appearance varied from clear to very muddy and analyses showed that the suspended matter ranged from 25 to 1000 ppm. At times, larger entrained bodies, such as leaves, twigs, and agglomerations of matter, could be observed passing through the test section. Turbidity and the larger solids had no measurable effect upon results. This corroborates the work of Randall (13) who reported a change of only 0.4 per cent in the speed of sound in water with 150,000 ppm of suspended matter.

Laboratory tests showed that turbidity caused by appreciable amounts of entrained air, visible as a white cloudiness, attenuated the ultrasonic signal an amount depending upon the frequency used and the transducer spacing. Such a degree of cloudiness was not observed during any tests utilizing water directly from the river. The introduction of large bubbles of air into the flowing water caused transients in the indicated water velocity.

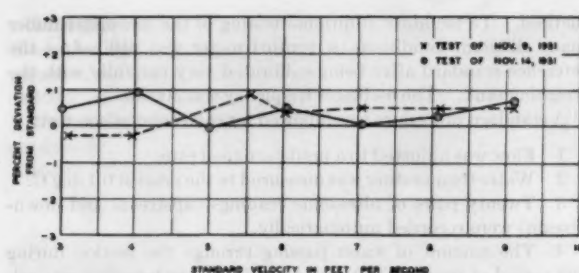
Temperature Effects. The propagation velocity of sound in water appears as a second-power quantity in Equation (3), the formula for calculating flow velocity. This velocity is 4865 fps at 20 C and increases approximately 9 fps per 1 deg C (14). In practice, only about 1 per cent error in the indicated flow velocity is introduced by an error of 3 deg C in the determination of water temperature in this range. As stated earlier, the water temperature was measured to 0.1 deg C during tests.

Long experience has shown that the temperature of the Susquehanna River does not change measurably over periods of several minutes' duration. Because of this, the relatively long switching cycle of 10 sec (15) was chosen to simplify instrumentation and to increase reliability and accuracy.

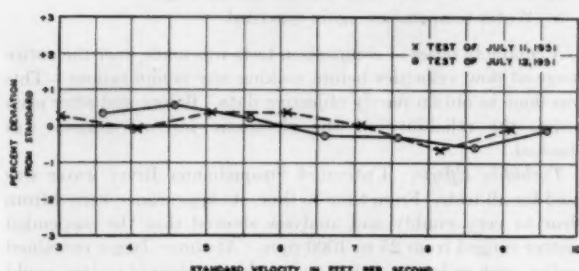
Effects of Nonuniform Flow. Although the measuring section of the ultrasonic flowmeter may be short, it has a finite length. Because of this, it is to be expected that asymmetric flow under certain conditions can introduce errors as a result of the change in flow pattern as the water passes through the measuring section. Analysis shows that velocity gradients through which the ultrasonic signal passes are seldom, if ever, great enough to introduce appreciable refraction and consequent error.

Fig. 3 shows the excellent precision of the flowmeter when the plane of asymmetric flow is parallel to the transducers. The data in this figure for the 12-in. transducer spacing are representative of those obtained from a series of nine tests made on as many different days. The data for the 6-in. transducer spacing are representative of four separate tests.

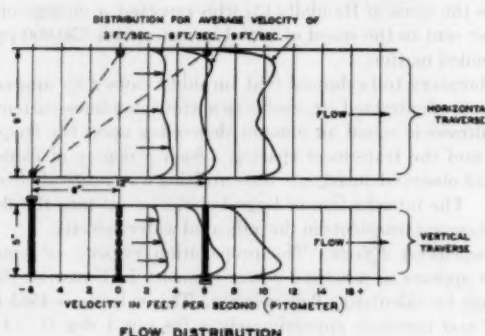
Fig. 4 indicates the magnitude of errors that can occur when



PRECISION OF MEASUREMENT WITH 6" TRANSDUCER SPACING



PRECISION OF MEASUREMENT WITH 12" TRANSDUCER SPACING

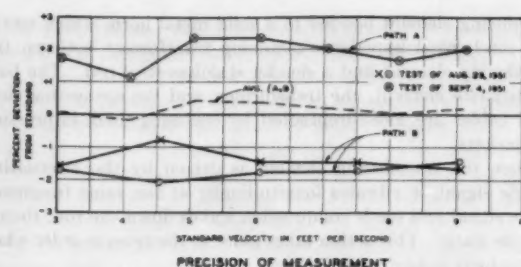


FLOW DISTRIBUTION

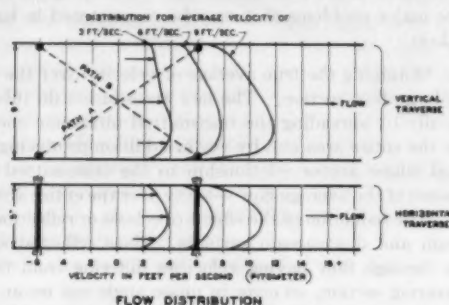
FIG. 3 TESTS WITH PLANE OF ASYMMETRIC FLOW PARALLEL TO TRANSDUCERS
(5-in. \times 9-in. duct; 6-in. and 12-in. transducer spacing.)

the plane of asymmetric flow is perpendicular to the transducers. Two independent sets of measurements were made, along path A and along path B, with the pair of transducers at the two different locations indicated. As the water passes through the measuring section, the higher velocities are reduced slightly and the lower velocities are increased correspondingly. When a given mass of water with a flow contour as shown in Fig. 4 reaches path A, the higher velocity is measured first. As this mass of water nears the second transducer, the lower velocity portion is measured. During the measuring interval, the lower velocity portion has been increasing in velocity. The measured velocity of the given mass of water is therefore higher than the true average value. When the measurement is made along path B, a correspondingly lower velocity is indicated. The graph in Fig. 4 shows these effects and how the average of the measurements along complementary paths A and B yields an accurate value of flow velocity.

Under the flow conditions of Fig. 4, the presence and direction of a transverse velocity were detected when the transducers were placed directly opposite each other, that is, when there was no transducer spacing on the principal axis of flow. The transverse



PRECISION OF MEASUREMENT



FLOW DISTRIBUTION

FIG. 4 TESTS WITH PLANE OF ASYMMETRIC FLOW PERPENDICULAR TO TRANSDUCERS
(5-in. \times 9-in. duct; 12-in. transducer spacing.)

velocities measured were of the proper magnitudes to account for the deviations observed. The exact nature of the phenomenon is not known; however, the use of vector analysis provides a plausible mathematical explanation. The tendency for the velocities to equalize in the measuring section produces a small vector of velocity, exactly opposite in direction to the pressure vector, across the pipe from the region of higher velocity toward the region of lower velocity. This vector has a component along the direct paths between transducers. Under the conditions shown in Fig. 4, the transverse velocity increases the flow velocity measured along path A and decreases it along path B.

Under the conditions of Fig. 3, the transverse vector of velocity is parallel to the transducers and perpendicular to the direct transmission path between transducers. Consequently, it has no component along the transmission path and introduces no measurable error. However, the transverse velocity theoretically can introduce an error of about one quarter of the maximum shown in Fig. 4. Under these circumstances, the path in the water would shift slightly and slightly change the length of the signal path in the transducer rods.

Turbulent nonuniform flow was produced by placing two or four uniformly spaced flat bars at the duct entrance. These bars were 5 in. high, 1 in. wide, and $1/4$ in. thick along the axis of flow. Exploration with a pitometer indicated the presence of extreme turbulence in the measuring section, although no effect was observed on the phase-angle recorder or on the cathode-ray oscilloscope used to observe instantaneous values of phase angle.

TESTS IN A TURBINE INTAKE, 16 FT \times 25 FT

The tests on the 5-in. \times 9-in. flow section demonstrated that the ultrasonic method is basically sound and warranted further development for use on a larger scale. A 42,500-hp adjustable blade-type turbine at the Safe Harbor plant was selected for the initial installation. The intake of the unit is divided into three separate bays and naturally would require three pairs of transducers to measure the total turbine discharge. Since precise current-meter testing equipment is available at Safe Harbor, it is possible to measure the discharge of a single bay; therefore, to re-

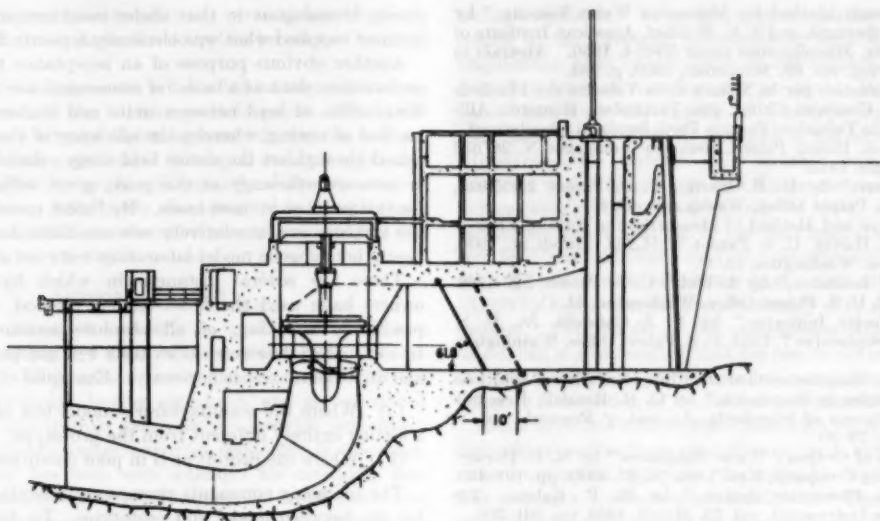


FIG. 5 LOCATION OF ANTENNA-TYPE TERMINATED TRANSDUCERS IN TURBINE INTAKE

duce the scope of the investigation, transducers were placed in the middle bay only. Furthermore, if river flow and available test personnel did not permit simultaneous comparison with current meters, a very good secondary standard is provided by the scroll-case piezometers and past data on distribution of discharge between the bays. The piezometers have shown remarkable consistency for more than 15 years of use as the metering elements for continuous measurement of discharge and hourly computation of plant efficiency. Previous field tests on five different units show that the maximum variation in the percentage of discharge through the middle bay is only 0.8 per cent.

Since the turbines are unwatered for maintenance only one time per year, considerable care was exercised in the design and installation of the underwater components to insure reliability. Waterproof housings were used for the electrical parts of the transducers and polyethylene tubing as conduits for the coaxial cables. Both the housing and tubing were pressurized with dry nitrogen. The $\frac{1}{4}$ -in.-diam transducer rods were recessed in the concrete walls of the intake by cutting an accurate groove with carborundum disks on a hand-power saw. The rods and the polyethylene tubing were fastened securely with clamps and anchor bolts. The clamps on the transducer rods were sound-insulated from the concrete and the clamps with rubber grommets; also the clamps were spaced at intervals which prevented resonant effects in the rods.

Fig. 5 is a cross-section view of the Safe Harbor turbine installation showing the location of the antenna-type transducers. The frequency used for tests in this unit was 25 kc.

Operating schedules did not permit simultaneous tests with current meters; however, a number of comparison tests were made with the piezometer metering as a secondary standard. Typical results are shown on the power-discharge curve in Fig. 6. The dashed-line curve through the piezometer points is about 0.5 per cent lower in discharge than the solid line through the ultrasonic points between approximately 6000 and 7000 cfs and, over the remainder of the test range, is as nearly the same as average curves can be drawn. It is recognized that these comparisons do not provide an absolute test of the precision of the ultrasonic flowmeter when used to measure large quantities of water. However, because of the long-term reliability of the piezometers used as the standard of comparison and the agreement between the

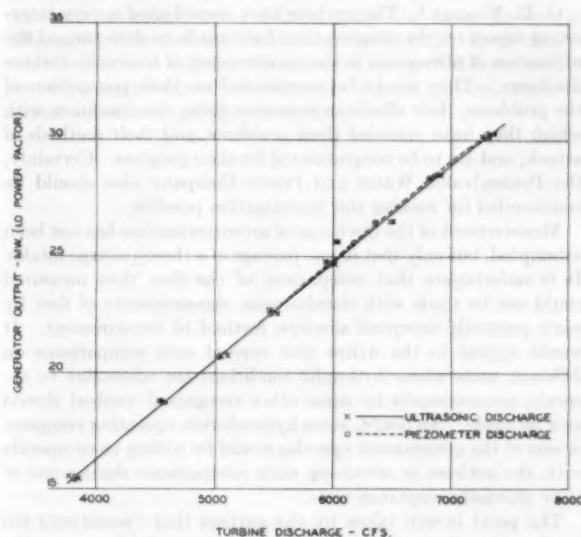


FIG. 6 COMPARISON TEST OF ULTRASONIC AND PIEZOMETER DISCHARGE METERS

discharges shown in Fig. 6, it is felt that the ultrasonic flowmeter will find definite application in flow measurements.

BIBLIOGRAPHY

- 1 "Ultrasonics and Their Scientific and Technical Applications," by Ludwig Bergmann, translated by H. S. Hatfield, John Wiley & Sons, Inc., New York, N. Y., 1943.
- 2 "Electromechanical Transducers and Wave Filters," by W. P. Mason, D. Van Nostrand Company, Inc., New York, N. Y., second edition, 1948.
- 3 "Ultrasonics," by Benson Carlin, McGraw-Hill Book Company, Inc., New York, N. Y., first edition, 1949.
- 4 "Piezoelectric Crystals and Their Application to Ultrasonics," by W. P. Mason, D. Van Nostrand Company, Inc., New York, N. Y., 1950, pp. 331-339.
- 5 "Ultrasonics," by P. Vigoreaux, John Wiley & Sons, Inc., New York, N. Y., 1951.
- 6 "Fluid Velocity Measuring Apparatus," by R. C. Swengel, U. S. patent applied for.

7 "An Ultrasonic Method for Measuring Water Velocity," by W. B. Hess, R. C. Swengel, and S. K. Waldorf, American Institute of Electrical Engineers, Miscellaneous paper 50-214, 1950. Abstract in *Electrical Engineering*, vol. 69, November, 1950, p. 953.

8 "Metodo Acustico per la Misura della Velocita dei Fluidi in Movimento entro Condotti Chiusi, con Particolare Riguardo All'Acqua Fluente Nelle Tubazioni Forzate Degli Impianti Idroelettrici," by Antonio Fiorensi, Italian Patent, Brevetto Industriale N. 287653 Rilasciato il 31 luglio 1931.

9 "Fluid Meter," by H. E. Hartig, U. S. Patent 2,015,933, Oct. 1, 1935, U. S. Patent Office, Washington, D. C.

10 "Fluid Meter and Method of Measuring the Rate of Flow of Fluids," by H. E. Hartig, U. S. Patent 2,151,203, March 21, 1939, U. S. Patent Office, Washington, D. C.

11 "Air Speed Indicator," by I. Wolff, U. S. Patent 2,274,262, February 24, 1942, U. S. Patent Office, Washington, D. C.

12 "Fluid Velocity Indicator," by N. J. Cafarelli, Jr., U. S. Patent 2,328,546, September 7, 1943, U. S. Patent Office, Washington, D. C.

13 "Ultrasonic Measurements of the Compressibility of Solutions and of Solid Particles in Suspension," by C. R. Randall, Research paper no. 402, Bureau of Standards, *Journal of Research*, vol. 8, January, 1932, pp. 79-99.

14 "Properties of Ordinary Water Substance," by N. E. Dorsey, Reinhold Publishing Company, New York, N. Y., 1940, pp. 191-193.

15 "Electronic Flowmeter System," by H. P. Kalmus, *The Review of Scientific Instruments*, vol. 25, March, 1954, pp. 201-206.

Discussion

G. H. VOADEN.⁵ The authors have contributed a very interesting report on the progress they have made to date toward the utilization of ultrasonics in the measurement of hydraulic turbine discharge. They are to be commended on their recognition of the problems, their efforts to overcome them, the frankness with which they have revealed these problems, and their methods of attack, and are to be congratulated for their progress. Certainly, the Pennsylvania Water and Power Company also should be commended for making this investigation possible.

Measurement of the discharge of an entire turbine has not been attempted, but only that in one passage of a three-passage intake. It is unfortunate that comparison of the flow thus measured could not be made with simultaneous measurements of flow by some presently accepted absolute method of measurement. It would appear to the writer that several such comparisons on different units where hydraulic conditions are conducive to accurate measurements by some other recognized method should now be made. No doubt, some hydroelectric operating company or one of the government agencies would be willing to co-operate with the authors in obtaining such comparisons during one of their planned acceptance tests.

The point is well taken by the authors that "sometimes too much emphasis is placed upon the results of field tests as the criterion of guaranteed performance rather than as a source of information for most efficient operation." The following further comments are offered in this connection:

If the turbine manufacturer has a laboratory-model test or a prototype field test of an homologous or nearly homologous unit, an absolute efficiency test is not necessary to determine whether or not the manufacturer has met his efficiency guarantees. Corrections can be made for the size differential between the model (or the previously tested homologous prototype) and the unit in question, and the efficiency of the latter can be calculated within plus or minus 1 per cent. Since the manufacturer usually has at least 2.0 per cent safety margin on his efficiency guarantees, it will be quite obvious whether or not there is any question as to the efficiency guarantees being met. Therefore this purpose of running an absolute efficiency test is fulfilled only when the manufacturer cannot produce a model or other test of a unit

closely homologous to that under consideration, or if a manufacturer supplied what was obviously a poorly finished runner.

Another obvious purpose of an acceptance test is to provide performance data as a basis for economical use of water flow and distribution of load between units and stations. The "index" method of testing, whereby the efficiency of the turbine is determined throughout the entire load range relative to a calculated or assumed efficiency at the peak, gives sufficient information for this purpose in most cases. By "most cases" is meant where the turbines are in relatively new condition and homologous or nearly homologous model-laboratory tests are available.

There are several instances in which hydroelectric-utility owners have used the index-testing method to the exclusion, practically speaking, of all absolute measurements of flow. In some cases, however, index tests will not provide the answer and an absolute test is necessary. Examples of such cases are:

- (a) Where the manufacturer's model test of the runner is in a setting entirely different from the prototype;
- (b) Where the prototype is in poor condition.

The foregoing comments prompt consideration of the step-up for size between model and prototype. To date no single scale formula has received world-wide acceptance although it is fully recognized that such a formula is theoretically justified, and comparisons of model and prototype performance in hundreds of instances over many years have removed all question that a differential exists. A tabulation of the many scale formulas proposed to date and a discussion of them may be found in a recent paper by S. P. Hutton.⁶

The scale formula which presently is in most common use in America is the Moody formula

$$E_p = 100 - (100 - E_m) \left(\frac{D_m}{D_p} \right)^n$$

where

E_p = efficiency of prototype turbine, per cent

E_m = efficiency of model turbine, per cent

D_m = diameter of model turbine

D_p = diameter of prototype turbine

n = an exponent, value of which is not definitely agreed upon, but usually a value of either 0.20 and 0.25 is used

The authors' point is well taken that "field testing has been limited by the short length of water passages and the geometry of the intakes." For many installations none of the accepted methods of absolute flow measurement is applicable, principally because of basic limitations of each of the several methods. Even under conditions of water passage which are ideal for the water-measurement method being employed, the field-test results are not necessarily accurate. The perfect field test has yet to be run. Where water-passage conditions are unfavorable for the method being employed, the field-test efficiency can readily be plus or minus 2 per cent or more from the actual.

The variation in efficiency of 2 per cent between individual units of identical design mentioned by the authors is rare. Perhaps the authors would elaborate somewhat on this point in their closure.

It is the writer's opinion that in the turbine proper, as produced by the manufacturer, any variations from the design can be controlled during manufacture so that resulting variations in efficiency between duplicate units will be well within the limits of accuracy of any field tests. For example, on recent index tests of two large "duplicate" Kaplan turbines, the shapes of the two

⁵ Assistant Chief Hydraulic Engineer, S. Morgan Smith Company, York, Pa. Mem. ASME.

⁶ Published by The Institute of Mechanical Engineers, 1 Birdcage Walk, Westminster, London, England, SW 1.

curves of power versus relative efficiency were almost identical, and one blade-gate cam shape was found to be correct for both units.

None of the above comments is intended as any adverse reflection on the work of the authors or their paper. There will always be uses for a simpler and more accurate way to measure flow of water. If such a method is developed and it is applicable to hydraulic turbines the writer will certainly reconsider some of his comments herein. Efforts are also being made to arrive at a better scale formula which can be more generally acceptable and this also would be very desirable.

AUTHORS' CLOSURE

Mr. Voaden's discussion presents an evaluation of indirect methods for obtaining the performance characteristics of hydraulic turbines. The authors are well acquainted with the index and scale formulas, having used both, and are convinced that data from these sources alone do not provide sufficiently accurate information to realize maximum economies in operation. The index method has been very valuable for extending performance characteristics over a wide range of head conditions after the index was proved reliable and its calibration determined from a direct method of field testing. The ultrasonic method in its present stage of development will provide reliable measurements in the type of intakes for which it was designed.

Throughout the years that hydraulic turbines have been used, a number of field tests have been made with other direct methods applicable to the installations. Although the results of these tests may be somewhat less than perfect, they do provide data believed to be more precise than that which can be obtained from

indirect methods. As an example, the authors have found the results of the two-type current-meter method very consistent. Duplication tests on the same 42,500 hp turbine have shown a deviation band of less than 0.5 per cent. Reports are also at hand that show the results of two-type current-meter tests conducted at two different locations in the intake of a 13,000 hp adjustable-blade turbine. The average velocity at the two locations varied 54 per cent and the geometric form was such to induce widely different angularity of flow, yet the efficiency agreement was as follows: 40 per cent load, 0.4 per cent; 70 per cent load (peak-efficiency load), 0.1 per cent; curves are coincident from 83 per cent to maximum load. One of the tests on the 13,000 hp unit was made simultaneously with another recognized method. The details of the comparison have not been published, but it is understood that the results are in close agreement. General knowledge of the detailed comparison should provide valuable information to the field of flow measurements.

Reliable results from turbine tests with current meters require meticulous calibrations, installation, and care of equipment. Aside from rather high costs, the method is limited by the type of flow passages and approach conditions. These considerations were the principal reasons for the development of the ultrasonic method.

The 2 per cent difference between supposedly identical units in the same plant was detected by the relative results of the two-type current-meter tests. Substantially the same percentage difference has been found in the maximum power capabilities of these machines. Perhaps the apparent rarity of differences of 2 per cent magnitude stems from the dearth of field tests on several turbines in multiple-unit plants.

Recent Investigations of the Mechanics of Cavitation and Cavitation Damage

By ROBERT T. KNAPP,¹ PASADENA, CALIF.

This paper describes water-tunnel investigations into the mechanics of "fixed"-type cavitation and into the probable mechanism through which this type causes material damage. High-speed motion pictures were used to study the cavity mechanics, and indications of the damage pattern were obtained by measuring the pitting rate on soft aluminum test specimens. Information was obtained on the frequency and intensity of the damaging blows, the distribution of damage in relation to the area covered by the cavitation, and the variation of the intensity of cavitation with velocity.

OBJECTIVE OF STUDY

THE purpose of this paper is to present additional experimental evidence on the mechanics of cavitation, with particular emphasis on those phases responsible for the production of damage on the adjacent solid guiding surface. In the past (1)² most of the work done in this field has been concentrated along two lines: (a) Investigations of the general mechanics of the cavitation process (2, 3, 4, 5, 6); (b) a study of the relative resistance of materials to cavitation damage (7, 8, 9). The objectives of most of the studies of the cavitation process have been the clarification of the behavior of the cavities in the liquid, and the development of analytical expressions to describe flow with cavitation. On the other hand, the investigations of the relative resistance of materials to cavitation damage have been concentrated largely on the behavior of the materials when subjected to some standardized type of cavitation. In addition to such empirical determinations of the relative resistance, considerable work has been done on the effects of the cavitation attack on the physical properties of the material. Thus the microstructure of the material has been studied to determine the physical changes produced by the cavitation which eventually lead to the removal of elements of the material from the surface. However, as yet there is only a partial understanding of the detailed mechanics by which the cavitation process in the liquid produces forces of sufficient intensity to cause physical damage on the guiding surface.

In 1917 Lord Rayleigh (10) presented a simple analysis of the collapse of an empty cavity in an ideal liquid which demonstrated that very high forces would be produced at the instant of collapse. However, it is only comparatively recent that the development of high-speed cinematography has made it possible to obtain experimental evidence showing that traveling cavitation voids collapse in good agreement with Rayleigh's analysis. This confirmation suggested the possibility of correlating the mechanics

of the cavitation process with the mechanics of surface damage. Study along these lines soon brought to light another difficulty. The agreement obtained between the Rayleigh analysis and the laboratory experiments was for individual traveling cavities. However, it developed that the more general type of cavitation occurring in hydraulic equipment is of the fixed-cavity variety. In this type of cavitation the main flow leaves the guiding surface and follows a free trajectory which usually returns to the surface at some downstream point. The fixed cavity is the space between the solid guiding surface and the free-liquid surface.

On preliminary examination this type of cavitation appears to be quite different from that caused by traveling cavities, since there seems to be no sudden collapse. If this were actually the case, the Rayleigh analysis would be inapplicable; thus some other mechanism would be required to explain the production of forces of high enough intensity to cause damage.

Since it is difficult to distinguish between the two types of cavitation with the unaided eye, there is only a little evidence to show which type occurs in given specific cases. However, the available evidence shows, reasonably conclusively, that damage is also produced by fixed-type cavitation. Hence it seemed necessary to make a more detailed investigation of the fixed-cavity mechanics to determine, if possible, the origin of the high-intensity forces that produce damage. This paper will be restricted to the discussion of laboratory experiments on fixed cavities and the evaluation of the results obtained.

DESCRIPTION OF EXPERIMENTS

Two types of experiments have been made. The first consists of a series of high-speed motion pictures to permit the study of the detailed mechanics of fixed-type cavitation on a simple body shape (11, 12). Various photographic reduction ratios and camera positions resulted in both general views of the entire cavitation region and detailed studies of the major features. Most of the work was done on a body of revolution consisting of a cylindrical center section, a hemispherical nose, and a long ogival afterbody. Attention was concentrated upon cavities which did not extend beyond the downstream end of the cylindrical section. These records showed that an irregular cyclic process was going on within the cavity, and enough indirect evidence was obtained to give a general idea of the nature of this process. However, it seemed desirable to confirm this with more direct evidence. Therefore a series of three geometrically similar two-dimensional bodies was constructed with semicircular noses, parallel mid-sections, and ogival afterbodies. Their mid-section thicknesses of 1, 2, and 4 in. were chosen to see if any significant differences in the flow patterns could be observed. Each had a span of 2 in., with end shields of clear plastic. High-speed motion pictures were taken, both in profile and in plan, for various cavity lengths at flow velocities of 25 and 50 fps. This was the highest velocity usable with the plastic end shields. However, velocities up to 100 fps were used with the bodies of revolution.

The second set of experiments was made on geometrically identical bodies of revolution, namely, a 2-in.-diam cylindrical mid-section with a hemispherical nose and an ogive afterbody. However, in this series, 3 in. of the cylindrical center section, on which the cavitation occurs, was constructed of a material of

¹ Professor of Hydraulic Engineering, Hydrodynamics Laboratory, California Institute of Technology. Life Mem. ASME.

² Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Hydraulic Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 17, 1954. Paper No. 54-A-106.

relatively low resistance to cavitation damage, since the major objective was the study of the development of damage and its relation to the physical characteristics of the cavitation. Two types of information were collected: (a) Relatively large-scale high-speed motion pictures of the cavitation actually occurring while damage was taking place; and (b) repetitive photomicrographs of representative areas along the length of the test section. These photomicrographs were closely spaced to cover the entire length of the cavity and also the adjacent downstream section. A technique was developed which made it possible to secure these photographs without removing the test body from the tunnel and without lowering the water surface enough to expose any portion of the body to the air. Thus the critical center sections were not handled during the entire run. As a matter of fact, they were not even touched. Therefore all changes in the successive photomicrographs were due to the cavitation.

Test Sections. The first test sections were made of free-machining brass which is known to have a relatively low resistance to cavitation damage, and is usually used as a calibration standard for the magnetostriction type of testing equipment. These test sections were run in the high-speed water tunnel for maximum periods of 25 hr at a flow velocity of 90 fps, and with cavity lengths of either 1 or 2 in. The photomicrographs showed definite evidence of damage, increasing with time. However, the damage rate was considered to be too low since the time required to test a single specimen was quite long and the tunnel operation, especially at this maximum permissible sustained speed, relatively expensive. It was felt that a satisfactory material should have the following properties: (a) It should show appreciable microscopic damage in an hour or less; (b) different lots obtained in the open market should all have the same physical properties; (c) it should be a metal, since most important cases of cavitation damage involve metallic surfaces; (d) its corrosion resistance should be relatively high compared to its cavitation resistance to minimize confusion between these two effects.

Commercially pure annealed aluminum conforms closely to these requirements. It is more readily damaged by cavitation than free-machining brass; it can be annealed readily in a simple furnace, after machining, to bring it to a standard and reproducible physical condition. Annealed 2S-F aluminum was therefore selected for trial. Since even before annealing, this material is very soft and therefore difficult to polish satisfactorily, a process was developed in the instrument shop which produced a satisfactorily smooth surface with a cutting tool. It was also found that the low temperatures required for subsequent full anneal did not change the surface finish.

Procedure for Typical Experiment. The procedure for a typical experiment was as follows: The test body was installed in the tunnel on a sting support and the tunnel was filled until the water was approximately 1 in. above the body. Then a microscope support and traverse mechanism were fastened to the open access window directly above the aluminum mid-section. A low-power microscope with camera attachment was used with a dipping cone so that the objective could be immersed to the proper depth. A Heine Ultropak illuminator was used, which is a variation of the vertical type. Photomicrographs were taken at a magnification of 30 diam with a field coverage of approximately 0.2 in. Twelve photographs were taken along the 3-in. test section with fields slightly overlapped to give complete coverage of the critical damage zone and 50 per cent coverage for the remainder of the section. Next, the tunnel was closed, filled, and the test run was made at the selected velocity, cavity length, and period. The tunnel was then stopped, the window opened, and another set of photomicrographs was taken of the identical areas.

Objectives of Test Runs. Since these runs were planned to supple-

ment the investigation of the fixed-cavity mechanics, the main objectives were to study the variation in damage along the length of the test section in relation to the position of the fixed cavity and to attempt to correlate the type of damage observed with the cavity mechanics. A further objective was to investigate the amount of damage as a function of the time of operation under cavitating conditions. The final objective was to attempt to secure preliminary information about the variation of the intensity of cavitation with changes in the velocity of the flow.

All tests except those used to explore the final objective were made with a tunnel velocity of 90 fps. Cavity lengths of 1 and 2 in. were studied, with separate aluminum test sections. A set of runs, starting with a period of 10 min and approximately doubling the length of each successive run, was made on a single test section until a total test period of 7½ hr had been obtained. Complete sets of 12 photomicrographs were taken (a) before the initial run, (b) between each successive run, (c) at the completion of the final test period.

During one run a high-speed motion picture was taken of the cavity at an exposure rate of 10,000 frames per sec. A similar set of runs was made on another aluminum test section for a cumulative time of 24 hr, and two otherwise identical sets were made using a 2-in. cavity length in place of the 1-in., the first one for a cumulative time of 12 hr and the second one for 40 min. This last set started with a 2½-min period because considerable pitting had been observed to take place in the initial 10-min period of the first three sets. A number of pits formed even in 2½ min; therefore another series of tests was run starting with a much shorter period. In this series a new test section was used for each run so that the damage on the entire section could be studied in more detail at a later time. It was found that the shortest practicable test period was 30 sec. Beginning with this, separate test sections were run for periods of 1¼, 2½, 5, 10, and 20 min.

The tests made to explore the final objective, i.e., the effect of variations in velocity on the intensity of cavitation, also employed separate test sections for each run. Since these tests were not planned until a preliminary analysis had been made of the results of the preceding studies, it was possible to take advantage of the experimental fact that for given conditions the rate of pitting did not vary with the length of the run. Thus the length of run was varied for the different velocities to obtain satisfactory counting densities of the pits. The runs were started at the maximum speed of the tunnel—100 fps; the next runs were made at 90 fps, and the third at 77½ fps. The run lengths were all 10 min. The rate of pitting observed on the test section from the latter run was so low that smaller steps in speed and longer times were selected for the subsequent tests, i.e., 71 and 63 fps, each for 20 min. The final run was made at 59 fps using three times the initial test period of 10 min. Under these conditions the pitting rate was so low that this preliminary series of tests was terminated.

EXPERIMENTAL RESULTS

Mechanics of Fixed-Type Cavitation. This discussion is based upon observations of cavitation on both the bodies of revolution and the two-dimensional bodies. All of these bodies had the same geometrical cross section, i.e., a semicircular nose, a parallel mid-section, and an ogive afterbody.

The physical concept of fixed cavitation is that such a cavity is formed by the breaking away of the flow from the guiding surface at a low-pressure point, usually assumed to be the point at which the pressure on the surface has fallen to the vapor pressure of the liquid. From this point the stream follows a trajectory determined by the pressure field, and usually recontacts the guiding surface at some distance downstream from the breakaway point. The fixed cavity is the space existing between the guiding surface and the free surface of the flow. The high-speed motion

pictures show that this cavity is fixed only in a statistical sense. Steady flow conditions are not established; instead, a continuous but somewhat irregular cyclical process develops. In each cycle three phases can be distinguished: (a) Formation and growth, (b) filling, and (c) breakoff.

For a given set of flow conditions, the maximum length of cycle occurs when it starts with no residual flow cavity, grows to the characteristic length, fills completely, and breaks away from the body. The breakoff time is negligible compared with that of the other phases. Shorter periods occur when breakoff comes after only partial filling, or when growth commences with a large residual cavity. Early breakoff seems to be associated with disturbances within the cavity rather than within the main flow. The stream of liquid that fills the cavity is the re-entrant flow which forms in the high-pressure area at the downstream end of the cavity. This re-entrant flow performs the necessary function of supplying the reaction force which turns the remainder of the flow downstream and parallel to the guiding surface. The upstream velocity of this re-entrant flow is proportional to the velocity at the free surface of the cavity and would be of equal magnitude if it were not for skin friction. However, since the re-entrant flow is a relatively thin sheet, skin friction is a large effect and rapidly reduces this upstream velocity.

Fig. 1 is a series of motion-picture frames showing the cavity growth, filling, and breakoff. The re-entrant flow is seen clearly in the fourth and fifth frame. The last frame shows the breakup of the surface which occurs at complete filling. As indicated in Fig. 2, the liquid in this re-entrant flow consists largely of the fluid that was adjacent to the free surface of the cavity. This is a significant factor in the explanation of the observed pattern of cavitation damage. Another significant detail shown in Fig. 2 is the stagnation zone which divides the re-entrant flow from the main flow. This is a zone of high relative pressure, since it is a full velocity head greater than the static pressure in the undisturbed flow.

The free surface of the cavity appears to be covered with a multitude of small cavities. Close inspection of the photographs shows that these small cavities first appear and grow rapidly to their ultimate size at the upper curved end of the main fixed cavity. After this initial growth, they move at substantially free-stream velocity and maintain a constant diameter until they reach the downstream end of the cavity, at which point they disappear.

If the behavior of the majority of these small traveling cavities is considered in conjunction with the cyclic process of cavity formation, filling, and breakoff, a rational picture develops. Most of these cavities, being small and apparently very close to the interface, will be in the re-entrant flow. The smaller ones will pass close to the stagnation point and larger ones may be swept into it. Some of the largest project into the main flow just outside of the stagnation streamline. However, all of these flow layers pass through the high-pressure area, and therefore the vast majority of these small entrained cavities must be collapsed completely. In the photographs a nebulous cloud flows away from the downstream end of the cavity during the filling part of the cycle. This may consist of minute traces of noncondensable

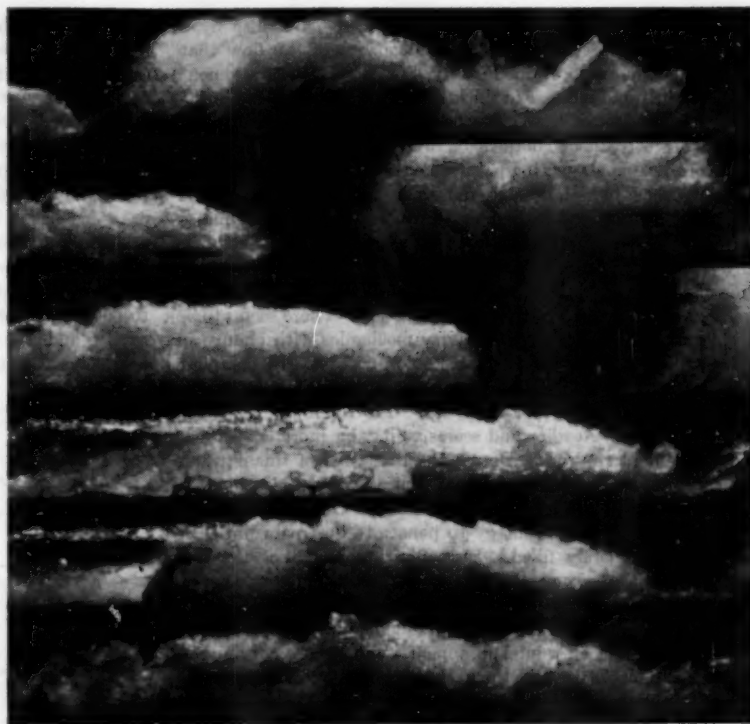


FIG. 1 CAVITY FORMATION AND RE-ENTRANT FLOW

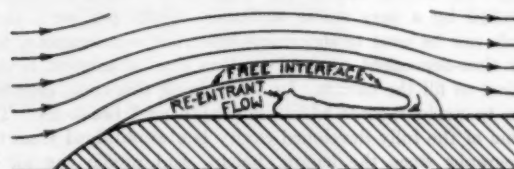


FIG. 2 FIXED-CAVITY FLOW

gas left after the condensation of the vapor which fills the voids.

Cause of Damage From Fixed-Type Cavitation. This description of the mechanics of fixed-type cavitation offers a simple explanation of the production of damage since the collapse of these small traveling cavities is concentrated into a narrow zone adjacent to the stagnation line. Furthermore, the pressures available to produce collapse are very much greater than the free-stream pressures; hence although the cavities are initially small, the final collapse velocities may well produce pressures high enough to exceed the yield point of the guiding surface and thus cause damage. In this type of flow the maximum pressure zone is on the surface and the direction of motion of the traveling cavities is toward the surface so that collapse may take place practically on the surface. Another important point is that, although the stagnation point remains relatively fixed at the end of the fully developed cavity for a major fraction of each cycle, nevertheless during the development phase, this stagnation point sweeps the entire length of the cavity. Owing to the many partial breakoffs that are observed, the relative time of dwell of the stagnation zone is a minimum at the upper end of the cavity and increases to a maximum at the equilibrium length. After breakoff the filled cavity acts as a foreign body over which the flow has to pass. It starts from rest and is accelerated to the stream velocity partly by the friction

force of the main stream on this slug of liquid, but also by a direct pressure force due to the impingement of the flow that passes over the newly formed cavity, Fig. 1. This impingement explains the rapid increase in thickness of the upstream end of the detached filled cavity. This process continues well below the normal downstream end of the fixed cavity, and since this impingement produces a stagnation point on or near the guiding surface, traveling zones of high pressure must sweep periodically along the body for a considerable distance downstream from the equilibrium length of cavity. Such high-pressure zones will collapse any traveling cavities which have escaped the first high-pressure area or which are rebounds of previously collapsed cavities.

Effect of Change of Velocity on Cavity Cycle. There seems to be a characteristic change in the cavity mechanics with change in velocity. The strong cyclical behavior with nearly complete filling of the cavity on each cycle seems to be associated with the higher velocities, i.e., from 50 fps up. For the lower velocities, the flow appears steadier and seems to have a more continuous entrainment at the rear. Detailed examination of this entrainment, however, shows that it is not uniform, but pulsating. It seems probable that for these lower velocities the re-entrant flow may not have sufficient energy to penetrate upstream until the cavity is completely filled. Instead, it may contact the free surface near the lower end of the cavity, and such contact may result in premature breakoff and regrowth. Since the amount of filling is small the corresponding frequency is relatively high, which makes the entrainment appear more nearly continuous. These last remarks describe the variation in behavior of cavitation on the 2-in.-diam models used in the water tunnel. It seems probable that for other sizes the velocity range for corresponding changes would vary since this behavior may be affected by both the body size and the properties of the liquid.

Since the taking rate of the high-speed motion pictures is controlled by a very stable oscillator, ± 0.02 per cent, these records can be used to determine velocities, accelerations, and other time-dependent quantities. One time of interest is that required to fill the cavity. In determining this time, care was taken to avoid records which showed partial breakoffs. The volume of the fully developed cavity is also obtained from the photographs since the length, height, and profile can all be measured. Knowing the cavity volume, the time between breakoffs, and the re-entrant velocity, it is possible to compute the thickness of the re-entrant flow. If a frictionless fluid is assumed, the speed of the re-entrant flow will be the same as that along the free interface, i.e., the stream velocity multiplied by $(1 + K)^{1/2}$. For the real fluid it is probable that the velocity of the re-entrant flow is more nearly that of the undisturbed stream.

It is also possible to estimate the thickness of the re-entrant flow from independent considerations. Gilbarg (13) has shown that when a cavity is formed downstream from a symmetrical curved obstacle, the drag on the obstacle is equal to the change in momentum of the liquid which forms the re-entrant jet. This relationship can be used to calculate the cross section of the re-entrant jet, since its velocity is known. In the experiments under discussion the cavity does not envelop the body, but collapses on the cylindrical center section. However, the basic nature of the phenomenon is unchanged. Thus it can be assumed that the increase in the nose drag caused by the presence of the cavity is equal to the change in the momentum of the re-entrant flow.

The experimental measurements of Rouse (14) have shown that upstream from the point at which the cavity springs clear, the pressure distribution on the nose of a body with cylindrical mid-section is unaffected by the presence of the cavity. These experiments cover a wide range of nose shapes from sharp-cornered flat noses to ogives, ellipsoids, and half-bodies. Hence

the increase in the nose drag due to the cavity can be calculated from the pressure distribution for noncavitating flow, since the cavity separates from the nose at the point where the ratio of the negative pressure to the velocity pressure equals the observed K for the given cavity lengths.

Fig. 3 shows the distribution of the axial component of the drag force on a hemispherical nose. This was plotted from the data of Rouse. The total area under the curve, multiplied by 2π , is the drag coefficient C_d . The shaded area is the increase in drag coefficient due to the presence of a cavity having a K of 0.378. This corresponds to one of the tunnel runs for which good high-speed motion pictures were available. These pictures were used to determine both the average longitudinal cross section of the cavity and the cycle time between major breakoffs. This completes the information required to estimate the re-entrant

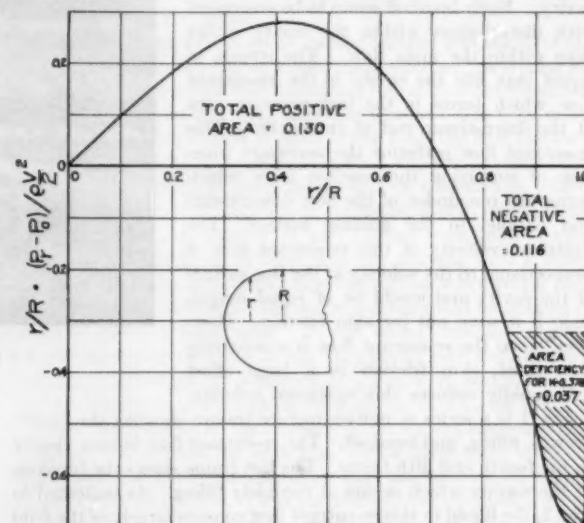


FIG. 3 AXIAL FORCE DISTRIBUTION ON HEMISPHERICAL NOSE

flow thickness from two independent points of view: (a) The cavity cross section and the observed filling time, and (b) the change in the momentum of the re-entrant flow which is equivalent to the increase in drag due to the presence of the cavity.

The momentum calculation is made as follows

$$\Delta C_d A \rho \frac{U^2}{2} = M(U + V)$$

in which

ΔC_d = change in drag coefficient

A = cross section of body = $\frac{\pi d^2}{4}$

ρ = mass density of liquid

U = velocity of undisturbed flow

M = mass rate of re-entrant flow = $(\pi d t)(\rho V)$

V = velocity of re-entrant flow = $U \sqrt{1 + K}$

D = diameter of body

t = thickness of re-entrant flow

This reduces to

$$t = \frac{\Delta C_d D}{8(1 + K + \sqrt{1 + K})} \cong \frac{\Delta C_d D}{8(2 + 1.5K)}$$

Table 1 gives a comparison of the re-entrant-flow thickness as calculated by the two methods for several different conditions.

In this table the agreement is surprisingly good, especially considering the difficulties involved in determining experimentally the cavity dimensions. It will be noted that with the exception of the first two examples which are for small-diameter bodies and short cavities, the thickness of the re-entrant flow, as calculated from the experiments, is somewhat greater than the flow predicted from the momentum theory. This is quite possibly due to the ejection of some of the liquid from within the cavity by chance contact of a portion of the re-entrant flow with the high-velocity interface. It also can be interpreted to mean that the

Graphical integration of the axial force on the upstream half of this body gives a drag coefficient of zero since the positive and negative zones agree in area within $1/2$ per cent. It is well known that the potential flow for a half-body has zero drag. In this connection Hoerner (16) states, "Half bodies represent the front end of sufficiently elongated bodies with parallel flanks. Positive as well as negative pressures are acting upon the frontal area and the resulting drag is zero, neglecting the small skin-friction forces. This phenomenon of zero drag holds theoretically true for all nose shapes." For this class of nose, the drag due to the presence

TABLE 1 COMPARISON OF RE-ENTRANT-FLOW THICKNESS CALCULATIONS, EXPERIMENTAL OBSERVATIONS VERSUS MOMENTUM THEORY PREDICTIONS

Body diam., in.	Velocity, fps	K	Cavity		Cycle time, sec	Re-entrant-flow thickness, in.		Thickness ratio, (1)/(2)
			Length, in.	Section, sq in.		Experimental (1)	Momentum theory (2)	
1/2	90	0.241	1.0	0.046	0.0058	0.0056	0.0099	0.67
1	90	0.348	1.0	0.0693	0.0050	0.011	0.0131	0.84
1	90	0.263	2.0	0.201	0.0088	0.0188	0.0182	1.03
2	90	0.409	1.9	0.197	0.0075	0.0205	0.0174	1.18
2	90	0.378	..	0.315	0.009	0.0276	0.0231	1.19
2	100	0.326	3.4	0.592	0.0122	0.0352	0.0297	1.18
2	90	0.322	3.6	0.532	0.0136	0.0315	0.0292	1.08
2	77 1/2	0.334	3.2	0.444	0.0146	0.0284	0.028	1.02
2	59	0.326	3.0	0.523	0.0198	0.0342	0.0297	1.15

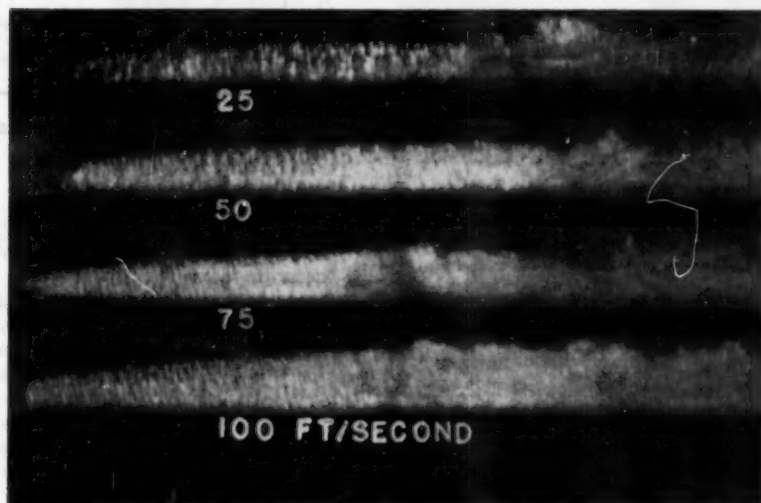


FIG. 4 EFFECT OF VELOCITY ON TRAVELING-CAVITY SIZE

cavity is not filled completely at breakoff so that the ratio of the two thicknesses indicates the degree of fill, the remainder being uncollapsed or rebound cavities, etc. Six of the nine runs used in the table were made with a tunnel velocity of 90 fps. It will be noted that the cycle time varies over a range of nearly 3:1. With the same diameter body it increases nearly linearly with the length of the cavity, as would be expected.

The last four runs are on the same body with approximately the same cavity length but with velocities varying from 100 down to 59 fps. Here, again, the variation in cycle time is rational as it varies approximately inversely with changes in speed.

It is interesting to note in Fig. 3 that the positive and negative pressure zones on the hemisphere differ in force by only 10 per cent. These are experimental measurements. Landweber (15) used as an example of his method of computing the potential flow about elongated bodies of revolution the rather blunt body whose equation is

$$y^2 = 0.04(1 - X^2)$$

of the cavity is the total form drag, just as for Gilbarg's completely enveloped body.

If high-speed photographs of fixed cavitation on the same body and same cavitation parameter are compared for a wide range of velocities, it is observed that the average size of the small traveling cavities decreases as the velocity increases. This can be seen in Fig. 4. A possible explanation is that the higher pressure gradients accompanying the higher velocities may produce additional cavities from weaker nuclei than were required for the lower velocities.

Mechanics of Damage. In this phase of the program an attempt has been made to correlate the development of cavitation damage in a flowing liquid with the mechanics of fixed-type cavitation. The damage pattern which developed on the annealed pure aluminum test sections furnished much information about the cavitation-force system and the reaction of the material to the attack. The hardness of the test specimens was very low, averaging 41 Vickers. The yield point is approximately 5000 psi. These low

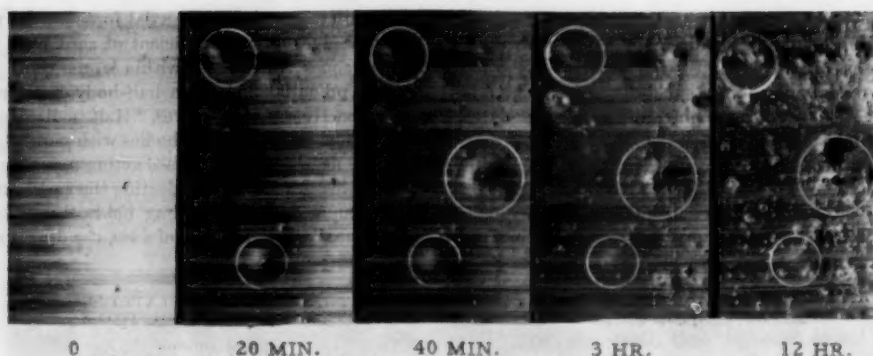


FIG. 5 DAMAGE DEVELOPMENT

values are desirable because all blows of sufficient intensity to produce physical damage of any type on the harder materials (bronze, stainless steel, etc.), used in hydraulic equipment where resistance to cavitation is required, will cause permanent deformation in the aluminum. Another useful property of pure aluminum is that the yield point does not change with rate of load application. Thus the same amount of energy will produce the same-size pit independent of the rate at which work is done. Fig. 5 shows damage development on a test section. The five 30-diam photomicrographs show the same area for increasing lengths of exposure. Several characteristic features may be observed:

1 The larger pits, such as those in the upper and lower white circles, show clearly that they have been formed by a single blow. Obviously, no metal has been removed, since the toolmarks can be seen, undiminished in sharpness, over the total area of the pit. Smaller pits appear to be formed in the same manner.

2 After 20 min, the pits are widely spaced with no other visible surface damage. This means that the cavitation developed only a few blows of damaging intensity during this long period.

3 If the micrographs for the longer periods are compared to the 20-min one, many of the original pits are found unchanged; the main difference is that new pits appear on the intervening surfaces. In some cases, e.g., in the large center circles, the pits have changed, apparently due to overlapping hits or "near misses."

Note that this damage was produced by cavitation in cold water at the relatively high velocity of 90 fps. This is well above the minimum velocity for which cavitation damage has been observed in many different metals in hydraulic equipment. Pure aluminum in the annealed condition has a very low hardness and yield point.

Frequency of Damaging Blows. The frequency of damaging blows is obtained by counting the number of pits, using a modified "blood-count" technique. The pits were classified roughly according to diameter: i.e., small, below 0.0025 in.; medium, 0.0025 to 0.005 in.; large, 0.005 to 0.01 in.; extra large, above 0.01 in.

Fig. 6 shows pitting rate versus exposure time for eight 1-in. cavity test sections and one 2-in. cavity. The numbers on the points are the test-specimen identification numbers. The velocity was constant. Although the range of exposure time is large, the number of pits appearing per unit time is about constant. It tends to decrease with increasing time, largely because duplicate hits escape detection. Motion pictures taken of the cavity surface during these runs were measured to determine the average size and "packing" of the traveling cavities. The cavities

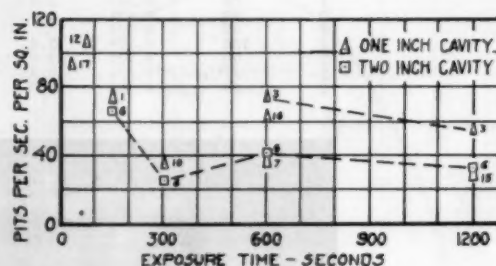


FIG. 6 EFFECT OF EXPOSURE TIME ON PITTING RATE

averaged 0.025 in. diam. Approximately 1.8×10^6 cavities per sec per in. of circumference were swept into the stagnation zone. Fig. 6 is for the maximum damage zone.

Figs. 9 and 10 show that the rate of pitting drops rapidly on both sides of the maximum. The summation over the 3-in. section is 60 pits per sec per in. of circumference. Hence only one in 30,000 of the traveling cavities causes a damaging blow. The question is—what happens to the other 29,999? The present guess is that the high-pressure zone above the surface acts as a sorting screen. The smaller cavities collapse so far away that the resulting shock wave is attenuated to below the damaging intensity before it reaches the surface. Everything favors the large cavities—higher collapse pressures with longer collapse times which carry them closer to the surface. This concept implies that liquid properties such as the size and number of the nuclei (17, 18) are of great importance, and may control the degree of damage for an otherwise constant cavitation condition.

In the zone of maximum damage an average of 56 pits, each with a mean area of 0.000003 sq. in., are formed per sec per sq. in. Thus each surface element is struck only once every 100 min or, assuming 24-hr operation, 5000 times per year. This low rate does not appear consistent with the commonly accepted concept that cavitation damage is essentially a fatigue failure caused by a large number of light blows. This force system seems more compatible with the concept that each blow produces permanent deformation and damage.

Intensity of Individual Cavitation Blows. It is possible, by using broad assumptions, to calculate the initial size of the collapsing cavity required to produce a damage pit. The Vickers diamond-pyramid hardness is obtained by measuring the dimensions of the indentation produced in the material by a known load. On the aluminum test sections the sizes of the hardness indentations and the damage pits are of about the same order, and the base-to-depth ratios are also comparable. The work done per unit



FIG. 7 PHOTOMICROGRAPHS OF PITTING WITH 1-IN. CAVITY



FIG. 8 PHOTOMICROGRAPHS OF PITTING WITH 2-IN. CAVITY

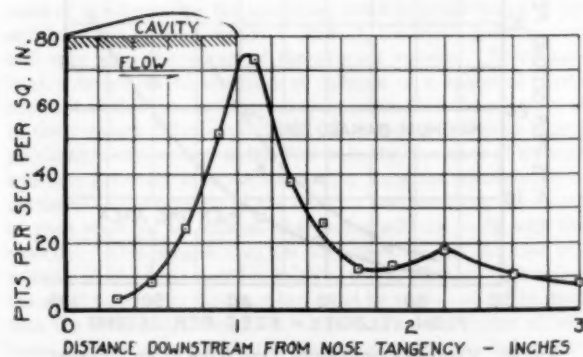


FIG. 9 DISTRIBUTION OF PITTING FOR 1-IN. CAVITY LENGTH

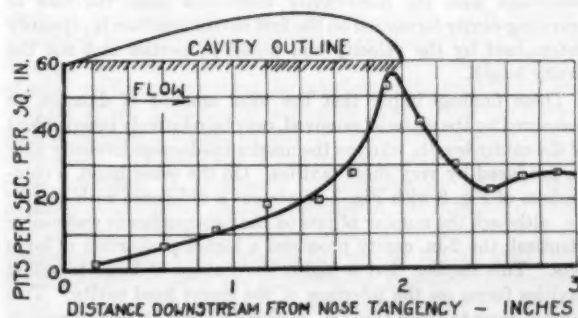


FIG. 10 DISTRIBUTION OF PITTING FOR 2-IN. CAVITY LENGTH

volume of indentation is easily calculated from the dimensions of the indentation and the known load applied by the hardness tester. It is about 60,000 in.-lb per cu in. Assuming the damage

pits are spherical segments, their volumes may be calculated from their diameters and depths. The work done in forming the pit is the product of its volume by 60,000, assuming that the work

of formation per unit volume is the same for the damage pit and for the hardness indentation.

The fraction of the total energy radiated from a collapsed cavity used in producing a pit may be assumed to be the ratio of the volume of the spherical pyramid whose base is the pit surface, to the volume of this sphere. Thus the required energy of the pit-forming cavity is obtained by dividing the energy used to produce the pit by this ratio. The diameter of this cavity is now calculable, as the work done (PV) by the fluid during collapse must equal the collapse energy. Since collapse must start as soon as the local pressure rises above vapor pressure, a reasonable estimate of the average pressure causing collapse is half the stagnation pressure. The results of such calculations for three typical damage pits are given in Table 2. These calculated cavities are considerably larger than the average cavities seen on the interface. This is consistent with the conclusions of the previous section.

Distribution of Damage. Fig. 7 shows strips from the 12 photomicrographs of a test section after a 10-min exposure to a 1-in. cavitation zone. If these strips were each rotated counter-clockwise 90 deg and placed end to end, they would be in their true relative position, with flow from left to right. Note the toolmarks.

TABLE 2 INITIAL CAVITY DIAMETER AND COLLAPSE CENTER LOCATION PREDICTED FROM DAMAGE-PIT DIMENSIONS

Damage pit				Collapsing cavity		
Diam. in.	Depth, in.	Work of formation, in.-lb	Volume ratio	Work of collapse, in.-lb	Initial diam. in.	Collapse center to surface, in.
0.005	0.0007	0.00042	0.077	0.006	0.073	0.005
0.007	0.0006	0.00052	0.029	0.018	0.106	0.010
0.0024	0.0004	0.000055	0.10	0.00054	0.033	0.002

In Fig. 9 the pitting rate is plotted against position along the test section. The relative size and location of the fully developed cavity is outlined at the top of the diagram. The maximum-damage zone coincides with the stagnation zone at the downstream end of this cavity. The pits above and below this zone presumably are formed during the sweeping of the stagnation zone along the full length of the test section during breakup. This distribution seems in good accord with the concept of fixed-cavity mechanics presented herein.

Figs. 8 and 10 are from a similar cavitation run for a 2-in. cavity. One striking feature is the lower pitting rate in the maximum-damage zone. However, the average pitting rate over the entire length of the test section is nearly identical for both cavity lengths, being 59 pits per sec for the 1-in. cavity and 60½ pits per sec for the 2-in. cavity. This agreement is also consistent with the fixed-cavity mechanics since the rate of traveling-cavity formation on the first cavity interface is primarily determined by the velocity and liquid properties and not the cavity length.

These findings imply that the total amount of damage, as measured by the material removed, may be relatively independent of the cavity length, whereas the maximum-damage intensity may be produced by very short cavities. On the other hand, a comparison of Fig. 8 with Fig. 7 points out a difference with length; i.e., although the number of pits on the two specimens was nearly identical, the 2-in. cavity produced a higher proportion of large pits. This implies that a higher percentage of large traveling cavities forms on the interface of the longer fixed cavity. Two factors seem to favor this, i.e., a slightly lower absolute pressure and a much longer time between inception and collapse. Probably the latter is the more important, for although, as previously stated, most of the growth in diameter of the traveling cavities takes place at the upstream end of the fixed cavity, there is some indication that slow growth continues until the high-pressure

zone is approached. Further investigation is needed to evaluate the relative importance to the production of damage of these opposing variables with cavity length, i.e., the lower pitting rate in the maximum damage zone versus the larger average size of pits.

Figs. 9 and 10 both indicate a small secondary maximum in pitting rate just downstream from the end of the cavity. The cause of this is not clear. One possible explanation, which has some confirmation from the photographic record, is that this increase may be due to the recollapse of some of the larger cavities which may rebound after first collapse in the stagnation zone.

Effect of Velocity on Cavitation Intensity. In the consideration of the results of this series of tests it must be remembered that it is a preliminary investigation consisting only of one small series of runs. Furthermore, it, like the rest of the experiments reported herein, applies only to fixed-type cavities. Finally, in the analysis of these tests, intensity of cavitation is assumed to be proportional to rate of pitting, modified by qualitative observations as to changes in the proportion of large pits produced.

Fig. 11 shows the observed pitting rate as a function of velocity. The circled points give the average rate of pitting over the entire length of the specimen. The crosses show the pitting rate in the

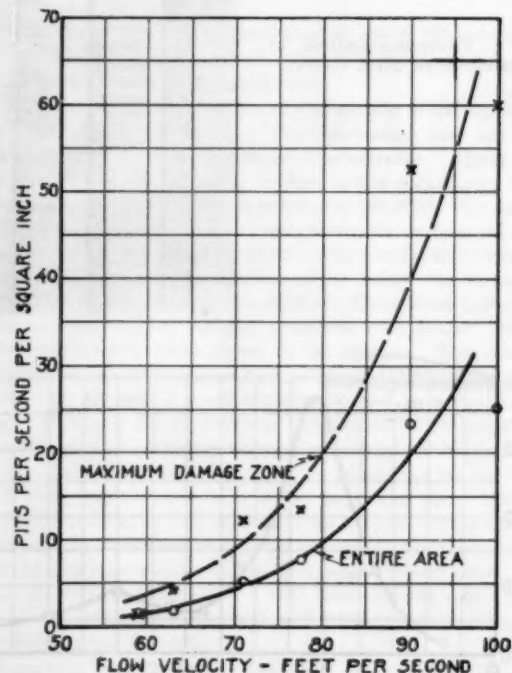


FIG. 11 VARIATION OF PITTING RATE WITH FLOW VELOCITY

maximum-damage area at the nominal end of the cavity. There is considerable scatter in the results. It is believed that this is due chiefly to variation in hardness between the test sections. A subsequent investigation showed that the time of anneal had not

been adequate to insure uniform properties of the different specimens. In spite of this scatter it is very clear that the rate of pitting falls very rapidly with decrease in speed.

Fig. 12 shows the same results plotted on logarithmic coordinates. For the 90 and 100-ft velocity rate the points plotted are the average of the different specimens tested. It will be seen that the points fall on the straight lines which best fit the two sets of points. Both have steep slopes, approximately six. This indicates that the pitting rate varies roughly with the 6th power

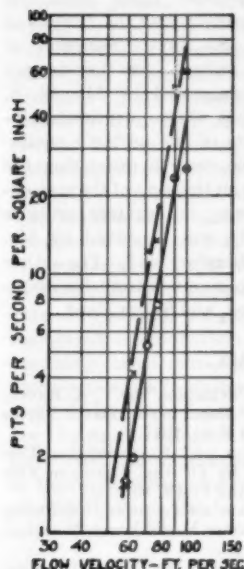


FIG. 12 LOGARITHMIC DIAGRAM OF PITTING RATE VERSUS FLOW VELOCITY

of the velocity. This is, of course, a tremendous rate of change. The evidence, however, is too meager to justify more than speculations concerning the implications.

Sixth-Power Variation of Pitting Rate With Velocity. The 6th-power variation implies that the pitting rate would about double for each $12\frac{1}{2}$ per cent increase in local velocity, assuming all other conditions, including size and location of the cavity, remain the same. Such a high sensitivity to the local velocity would help to explain some of the apparent discrepancies in the amount of cavitation damage observed on similar machines operating under nearly the same field conditions. Another type of implication carried by these results is the great importance of using a carefully standardized and controlled cavitation intensity in the testing of the relative resistance of materials.

The reason for this high sensitivity of the cavitation intensity to change in velocity is not obvious. Since the cavity size remained constant throughout these experiments, the geometry of the flow was unchanged. Hence the velocity distributions at any cross section were similar for all speeds. However, the base pressure, i.e., the cavity pressure, has the same absolute value for all the runs, i.e., the vapor pressure of the liquid. The maximum pressure rise above this base was equal to the velocity head of the flow along the interface, which is proportional to the square of the average velocity. In fact, the entire pressure field will vary with the square of the average velocity. Now, Rayleigh's derivation of the time of collapse of a spherical cavity shows that this time is inversely proportional to the square root of the pressure difference. Since this pressure difference is proportional to the square of the flow velocity, the time of collapse will vary inversely with the velocity, or since the distance is proportional to the product of the velocity and the time, the length of path required to collapse the cavity will not vary with the velocity. This means that for all velocities the cavities will collapse in the same position relative to the surface. Therefore the great increase in pitting rate cannot be explained on the basis that the average distance from the surface to the center of collapse would be less with the higher velocity. An increase in pitting rate will be anticipated if the average cavity size increased with velocity. Fig. 4, however, shows that, in general, the average size of the cavity appears to decrease slightly with increase in velocity. This should tend to decrease rather than increase the pitting rate.

Attempts to measure the average size of cavity from these

pictures show that the variations with velocity are too small to be distinguishable within the available definition. Thus it can be said that, within the accuracies of these experiments, the cavity size remains sensibly unchanged with changes in velocity. Since the energy of collapse is the product of the initial volume of the cavity and the pressure causing collapse, this energy will increase with the square of the velocity. Therefore, as the velocity increases, a given-size cavity collapsing at a given point will produce a larger pit and also cavities collapsing at a greater distance from the surface will be able to produce pits.

Another factor which should increase the rate of pitting is that the number of cavities swept into the collapse zone increases linearly with the velocity. At first sight it would appear that the combination of these two factors should increase in pitting rate with the cube rather than the 6th power of the velocity. However, since the number of cavities swept into the collapse zone for each pit is so great, it appears certain that a very small increase in the radius of effective action of the collapsing cavities would greatly increase the number of cavities able to produce pits. At the present time these three factors seem to offer the most acceptable explanation of the high sensitivity of the rate of pitting to change in velocity.

Pit Size-Distribution. In the discussion of the method of counting the pits, it was explained that, during the counting, the pits were classified according to diameter into four groups, with the three larger groups each covering a range of 2:1 in diameter and with the smallest group consisting of all those pits below 0.0025 in. diam. On the previous runs, which were all at 90 fps, it was observed that over 95 per cent of all pits were in the smallest group. An examination of the counts for the speed runs shows that the percentage of pits in the three larger classes varies with velocity as follows:

Ft per sec	Per cent
100	5.1
90	4.1
77 1/2	2.6
71	2.65
63	0
59	0

These results do not justify plotting, but it appears probable that the percentage of large pits also increases with some large power of the velocity; hence it may well be that the damage rate, as measured by rate of removal of material, will be found to vary even more rapidly than the 6th power of the velocity.

SUMMARY AND CONCLUSIONS

Fixed-type cavitation appears to be one of the most prevalent forms of cavitation occurring in flowing liquids. The idealized concept of fixed cavitation includes no mechanism for producing damage. A study of the real phenomenon shows that a multitude of small cavities sweep along the interface of the fixed cavity and collapse in the downstream stagnation zone. Thus the mechanism of damage is similar to that for the simpler traveling type of cavitation. The small size of the traveling cavities in the fixed-type cavitation is apparently counterbalanced by the high pressure available to produce collapse in the stagnation zone. One important feature of the fixed cavity is re-entrant flow at the downstream end. This upstream flow rapidly fills the cavity and is the primary cause of the characteristic growth-filling-breakoff cycle. Because of this cyclic process the stagnation zone sweeps the entire cavity length and the downstream region as well, thus creating relatively large areas of moderate cavitation damage on both sides of the maximum-damage zone.

Much significant information about the type and intensity of the cavitation attack can be derived from the pitting pattern on soft metallic specimens. One surprising fact is the relatively low

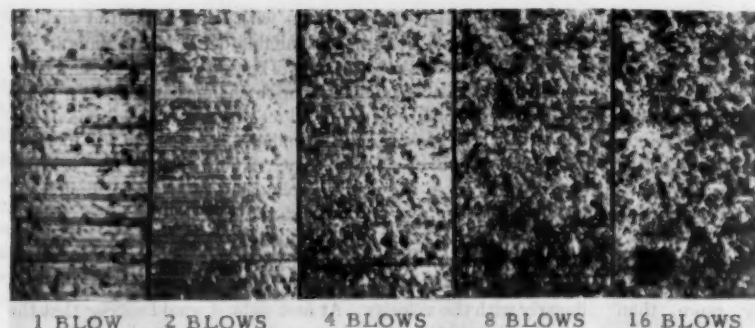


FIG. 13 DAMAGE DEVELOPMENT IN MAXIMUM DAMAGE ZONE

pitting rate. Thus the annealed aluminum used in this study is soft enough to pit under an applied force of much lower intensity than the fatigue limits of the materials used in the construction of hydraulic equipment. Nevertheless, at 90 fps, which is well above the lower limit for cavitation damage on these materials, a given point on the surface is hit only once every 100 min. This leads to the tentative conclusion that damage produced by fixed-type cavitation is caused by relatively infrequent high-intensity blows which either remove particles of material by each blow or by early fatigue. It will be recalled that, on the basis of the pitting rate and the average diameter of the pits, it was calculated that each point on the surface would, on the average, be struck once every 100 min. By chance, the time intervals used in one damage run have a 90-min base, corresponding roughly to the average time between blows. Thus the five photomicrographs in Fig. 13 show the "blow-by-blow" damage to a typical area in the maximum-damage zone.

In the past a basic obstacle to the study of the mechanics of cavitation damage has been the lack of any measure of intensity of cavitation. The discovery that fixed cavitation produces damage blow by blow, and that the repetitive rate is relatively low, suggests the use of a soft metal brought to a standardized physical condition for a working measure of cavitation intensity. For this use the specimen should be exposed to the cavitation for only a few minutes to avoid superposition of pits. The determination would be made by measuring the size range and rate of formation of the pits in place of the usual loss-of-weight measurement. The damage-resistant properties of the soft test specimen are not relevant. This material would only be used to obtain a record of the intensity and frequency of all of the blows capable of damaging, in any manner, the metal structure of the equipment under consideration.

One significant example of information obtainable by this approach is the fact that the total number of damaging blows struck in a cavitating region in a given time seems primarily a function of the physical properties of the liquid and the relative velocity of flow. Reasonable changes in the cavity length, with corresponding changes in the cavitation parameter, do not alter the total number of blows struck. Thus further study may show that total damage is independent of the length of the cavitation zone and intensity of damage may be greatest for short cavities.

Additional work needs to be done to clarify the effect of velocity variation on pitting rate and also to correlate pitting rate and pit size on standardized test specimens with damage as measured by rate of removal of material. Also, it begins to appear important to investigate the variation in intensity of cavitation with the size and shape of the cavity. Preparations are now under way to study pitting rate on a series of geometrically similar bodies of the same general shape as those reported herein, but varying in size over a relatively wide range, and also to in-

vestigate the pitting rate on bodies of the same size but with different nose shapes to see the effect of this variable and whether it can be isolated from that of velocity of flow and cavity length.

ACKNOWLEDGMENTS

The primary support for the work described herein was received from Contract Nonr-220(08), which contains funds from the Office of Naval Research and the Bureau of Ships of the U. S. Navy. Selection of aluminum as the material for the damage test sections was suggested by Mr. W. S. Rheingans, who has shown continuing interest in all aspects of the author's cavi-

tion research program. The principle used in estimating the stream thickness of the re-entrant flow on the basis of the momentum required to balance the deficiency in negative pressure caused by the presence of the cavity was suggested by Mr. Joseph Levy of the Hydrodynamics Laboratory staff. The author also gratefully acknowledges the help received from his assistants, Frank Bonamassa, James Isaacson, and Salyer Reed.

BIBLIOGRAPHY

- 1 "An Annotated Bibliography of Cavitation," by F. A. Raven, A. M. Feiler, and Anna Jespersen, Navy Department, David Taylor Model Basin, Washington, D. C., Report R-81, 1947.
- 2 "Untersuchungen über Kavitation und Korrosion bei Turbinen, Turbopumpen und Propellern," by Dr.-Ing. Hermann Föttinger, "Hydraulische Probleme," 1926, pp. 14-64, 107-110.
- 3 "On the Mechanism and Prevention of Cavitation," by Phillip Eisenberg, Navy Department, David Taylor Model Basin, Washington, D. C., Report 712, 1950.
- 4 "A Brief Survey on the Progress of the Mechanics of Cavitation," by Phillip Eisenberg, Navy Department, David Taylor Model Basin, Washington, D. C., Report 842, 1953.
- 5 "Water Tunnel Investigations of Steady State Cavities," by H. L. Pond and Phillip Eisenberg, Navy Department, David Taylor Model Basin, Washington, D. C., Report 668, 1948.
- 6 "Cavitation Characteristics and Infinite-Aspect-Ratio Characteristics for a Hydrofoil Section," by J. W. Daily, Trans. ASME, vol. 71, 1949, p. 269.
- 7 "Kavitation und Kavitationskorrosion," by J. Akeret, "Hydromechanische Probleme des Schiffsantriebs," 1932, p. 227.
- 8 "Werkstoffzerstörung durch Kavitation," by Hans Nowotny, VDI-Verlag, Berlin, Germany, 1942, 84 pp.
- 9 "Accelerated Cavitation Erosion," by W. S. Rheingans, Trans. ASME, vol. 73, 1950, p. 705.
- 10 "On the Pressure Developed in a Liquid During the Collapse of a Spherical Cavity," by Lord Rayleigh, *Philosophical Magazine*, series 4, vol. 34, 1917, pp. 94-98.
- 11 "Laboratory Investigations of the Mechanism of Cavitation," by R. T. Knapp and A. Hollander, Trans. ASME, vol. 70, 1948, p. 419.
- 12 "Cavitation Mechanics and Its Relation to the Design of Hydraulic Equipment," by R. T. Knapp, James Clayton Lecture, Proceedings of The Institution of Mechanical Engineers, series A, vol. 166, 1952, pp. 150-163.
- 13 "Free Boundaries and Jets in the Theory of Cavitation," by David Gilbarg and James Serrin, *Journal of Mathematics and Physics*, vol. 29, April, 1950, pp. 1-12.
- 14 "Cavitation and Pressure Distribution," by Hunter Rouse and J. S. McNown, Bull. 32, No. 42, State University of Iowa, 1948.
- 15 "The Axially Symmetric Potential Flow About Elongated Bodies of Revolution," by L. Landweber, Report 761, David Taylor Model Basin, Navy Department, Washington, D. C., August, 1951.
- 16 "Aerodynamic Drag," by S. F. Hoerner, 1951, published by the author, 148 Busted, Midland Park, N. J.
- 17 "Cavity Formation in Water," by E. N. Harvey, W. D. McElroy, and A. H. Whiteley, *Journal of Applied Physics*, vol. 18, February, 1947, p. 162.
- 18 "The Effect of Air Content on the Appearance of Cavitation in Distilled, Salt, and Sea Water," by F. Numachi and T. Kurokawa, Translation and Commentary, Ordnance Research Laboratory, Pennsylvania State College State College, Pa., 1946.

On the Mechanism of Cavitation Damage

By M. S. PLESSET¹ AND A. T. ELLIS,² PASADENA, CALIF.

A new method for producing cavitation damage in the laboratory is described in which the test specimen has no mechanical accelerations applied to it in contrast with the conventional magnetostriction device. Alternating pressures are generated in the water over the specimen by exciting a resonance in the "water cavity." By this means the effects of cavitation have been studied for a variety of materials. Photomicrographs have been taken of several ordinary (polycrystalline) specimens and also of zinc monocrystals. The zinc monocrystal has been exposed to cavitation damage on its basal plane and also on its twinning plane. X-ray analyses have been made of polycrystalline specimens with various exposures to cavitation. The results show that plastic deformation occurs in the specimens so that the damage results from cold-work of the material which leads to fatigue and failure. A variety of materials has been exposed to intense cavitation for extended periods to get a relative determination of their resistance to cavitation damage. It is found that, roughly speaking, hard materials of high tensile strengths are the most resistant to damage. While this survey is not complete, it has been found that titanium 150-A and tungsten are the most resistant to damage of the materials tested. Cavitation-damage studies, which have been carried out in liquid toluene and in a helium atmosphere, show that chemical effects can be, at most, of secondary significance.

INTRODUCTION

NUMEROUS studies have been made of the cavitation process and the mechanism of cavitation damage. No attempt will be made here to give a bibliography on this subject; reference may be made, however, to the monograph by Nowotny³ and to an investigation of damage with a magnetostriction oscillator by Kornfeld and Suvorov.⁴ In laboratory studies of cavitation damage it is often desirable to accelerate the damage rate. It has long been recognized that the cavitation produced by pressure variations at relatively high frequencies will produce damage more rapidly than the cavitation under ordinary flow conditions. The procedure which has been universally used consists in high-frequency acceleration of the test specimen relative to the liquid medium. These accelerations are obtained by application of the magnetostrictive effect in a nickel rod to the end of which the test specimen is attached. The liquid partakes of this accelerating motion and heavy cavitation is produced.^{5, 6} In such experiments the quantities measured are the displacement of the test surface and its frequency. Weight loss of the specimen also is determined as a function of exposure time to the cavitation field.

¹ Professor of Applied Mechanics, California Institute of Technology. Mem. ASME.

² Senior Research Fellow, California Institute of Technology.

³ "Werkstoffzerstörung durch Kavitation," by H. Nowotny, Edward Brothers, Inc., Ann Arbor, Mich., 1946.

⁴ "On the Destructive Action of Cavitation," by M. Kornfeld and L. Suvorov, *Journal of Applied Physics*, vol. 15, 1944, p. 495.

Contributed by the Hydraulic Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 11, 1954. Paper No. 54-A-76.

A new method for generating cavitation damage has been developed and studied by one of the authors and appears to have some important advantages for this type of investigation. While this method has been described,⁵ it is necessary to give a brief description of it here so that the results of the present study may be more easily interpreted.

PRESSURE VARIATIONS PRODUCED IN A RESONANT CAVITY

The apparatus used consists of a cylindrical beaker containing water with a barium titanate ring just below the surface, Fig. 1. If an alternating electric field is applied across the conducting coatings on the inner and outer surfaces of the ring, the volume of the ring oscillates with the applied electric field. At the proper

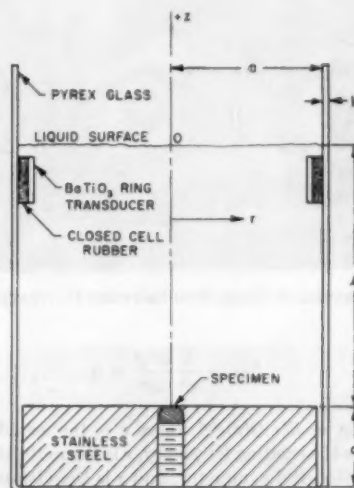


FIG. 1 DIAGRAM OF CYLINDRICAL BEAKER AND CO-ORDINATE SYSTEM

frequency of the electric field, a standing wave pattern is set up in such a way that a large pressure amplitude is produced at the center of the bottom plate. The pressure variation is sufficient to produce cavitation over the specimen with a sinusoidal voltage of amplitude between 100 and 200 volts. For some of the experiments, the oscillation frequency was 18,000 cycles; for the remainder of the studies, with a beaker of different dimensions, the oscillation frequency was 24,000 cycles. Fig. 2 shows the well-defined cavitation cloud over the surface of the specimen. Fig. 3 shows the complete apparatus with the audio oscillator, the power amplifier, and the beaker.

The acoustic theory for the pressure field in a cylindrical cavity is easily developed. The fluid velocity \vec{v} and the time varying part of the pressure p are related to the velocity potential ϕ by $\vec{v} = -\nabla\phi$ and $p = \rho_w \partial\phi/\partial t$, where ρ_w is the density of water. Viscous and streaming effects are neglected. One seeks a characteristic solution of the wave equation

⁵ "Production of Accelerated Cavitation Damage by an Acoustic Field in a Cylindrical Cavity," by A. T. Ellis, Hydrodynamics Laboratory Report No. 21-14, California Institute of Technology, Pasadena, Calif.

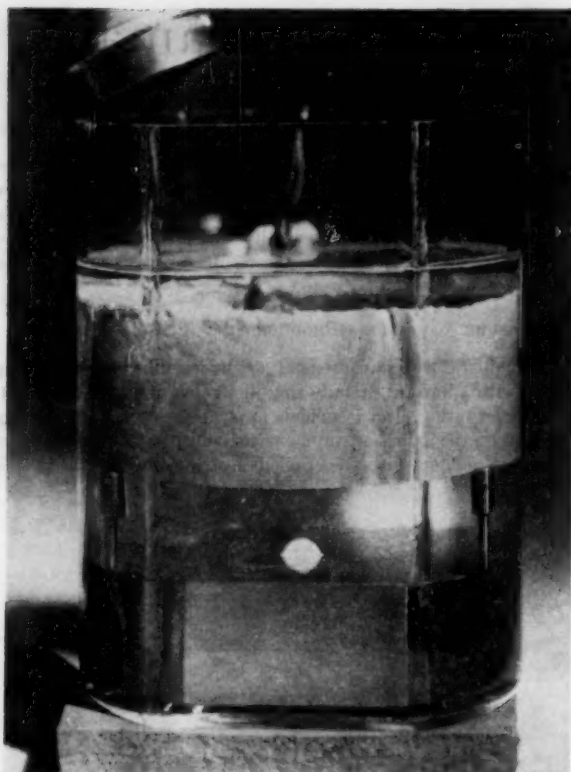


FIG. 2 BEAKER IN OPERATION SHOWING CAVITATION CLOUD

$$\Delta \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = 0 \dots \dots \dots [1]$$

of the form $\varphi = R(r)Q(z)e^{i\omega t}$ where c is the sound velocity in water, $\omega/2\pi$ is the exciting frequency, r is the radial distance from the axis of cylindrical symmetry, and z is the distance measured along this axis, Fig. 1. The origin of z is at the water surface and the positive direction of z is upward. The boundary conditions may be described as follows: The massive steel plate at the bottom of the container is assumed to move as a free rigid mass. Under this assumption, Newton's second law of motion gives a condition on the velocity potential at the interface of the liquid and the steel plate

$$\rho_s \frac{\partial \varphi}{\partial t} - \frac{d}{dt} \left(\frac{\partial \varphi}{\partial z} \right) (\rho_s d) = 0, \text{ at } z = -l$$

or

$$\varphi - \frac{\partial \varphi}{\partial z} \left(\frac{\rho_s d}{\rho_w} \right) = 0, \text{ at } z = -l \dots \dots \dots [2]$$

where ρ_s is the density of the steel plate and d is its thickness. The boundary condition at the beaker wall is determined under the assumption that this wall is under simple hoop tension. The force per unit length of the cylindrical wall is $Ebda/a$ where E is the Young's modulus of the glass (pyrex), b is its thickness, and a is its unstrained radius. Then the force equation at the wall, $r = a$, leads to the boundary condition

$$\varphi - \left(\frac{E}{a^2 \omega^2} - \rho_s \right) \frac{b}{\rho_w} \frac{\partial \varphi}{\partial r} = 0, \text{ at } r = a \dots \dots \dots [3]$$

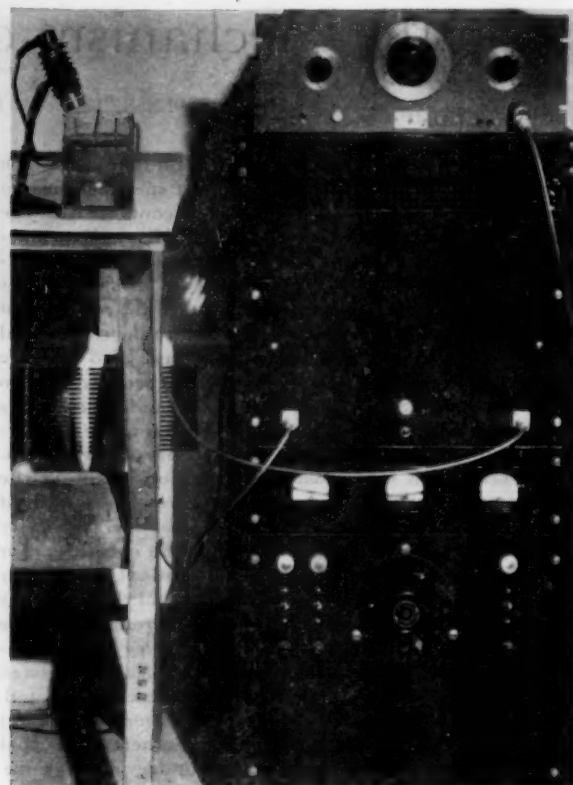


FIG. 3 COMPLETE APPARATUS INCLUDING OSCILLATOR, AMPLIFIER, AND BEAKER

where ρ_s is the density of the glass. There is finally the boundary condition at the free surface of the water

$$\varphi = 0, \text{ at } z = 0 \dots \dots \dots [4]$$

The characteristic solutions which satisfy these boundary conditions are

$$\varphi_{mn} = A_{mn} \sin \beta_{mn} z J_0 \left[\left(\frac{\omega_{mn}^2}{c^2} - \beta_{mn}^2 \right)^{1/2} r \right] \dots \dots \dots [5]$$

where J_0 is the Bessel function of zero order and β_{mn} is a solution of the equation

$$\frac{\tan \beta l}{\beta} = - \frac{\rho_s d}{\rho_w}$$

The angular frequency ω_{mn} is given by

$$\omega_{mn} = c \left[\left(\frac{\alpha_{mn}}{a} \right)^2 + \beta_{mn}^2 \right]^{1/2}$$

where the constant α_{mn} is determined from the solution of the equation

$$\frac{J_0(\alpha)}{\alpha J_1(\alpha)} = \frac{b}{a} \left[\frac{\rho_s}{\rho_w} - \frac{E}{\rho_w c^2 (\alpha^2 + a^2 \beta_{mn}^2)} \right]$$

A series of characteristic frequencies has been computed and has been compared with measured resonant frequencies for the water cavity. Agreement between theoretical and observed values is very good. A comparison of theoretical and ob-

TABLE 1 THEORETICAL AND OBSERVED CHARACTERISTIC ANGULAR FREQUENCIES

Dimensions of the pyrex beaker are $a = 4.83$ cm, $b = 0.254$ cm, $d = 2.54$ cm, $l = 8.13$ cm. Young's modulus for pyrex $E = 6.76 \times 10^{11}$ dynes/cm²; velocity of sound in water at room temperature $c = 1.52 \times 10^4$ cm/sec; density of water $\rho_w = 1$ gm/cc, density of stainless-steel bottom plate $\rho_s = 7.93$ gm/cc, density of pyrex $\rho_p = 2.23$ gm/cc.

ω/mm	Calculated values	Measured values
0011	88	85
0001	117	117
0001	168	167
0011	217	213
0012	167	171
0018	253	
0014	340	
0022	184	

served characteristic frequencies for specific beaker dimensions is shown in Table 1. The pressure amplitude has been measured as a function of r and z . These variations also agree well with those determined by Equation [5]. Of immediate practical importance is the fact that the most easily excited mode has the largest pressure amplitude at the center of the bottom plate. Thus the cavitation cloud is readily confined to the small region over the specimen. Fig. 4 shows the cavitation bubble cloud at two different points of the pressure cycle. The upper picture shows the bubble cloud near the pressure maximum and the lower picture shows the bubble cloud near the pressure minimum. Fig. 5 gives somewhat greater detail in the bubble cloud near the maximum (the lower half of the picture is a reflection of the bubble cloud in the stainless-steel plate).

In the opinion of the authors, the method used here for producing cavitation damage in the laboratory has several important advantages over the magnetostriction oscillator method. The use of the barium titanate electric transducer to generate standing pressure waves in the water cavity results in an apparatus which is very simple electrically and which, at the same time, is appreciably less expensive than the magnetostriction oscillator. The present device is stable in operation and requires a minimum of maintenance and attention. A wide range of frequencies is readily available. Of perhaps greater significance is the fact that one has here an efficient way of producing the oscillating cavitation field. As a consequence, there is negligible heating of the water or the test specimen even in very long exposures at high powers. Under heavy cavitation conditions the water temperature has been observed to rise less than 3 deg C in 1 hr. It also should be remarked that there is negligible streaming of the water in the container and there is no pumping action on the water. An additional consideration which led the authors to the adoption of this resonant "cavity" procedure was the fact that the test specimen is stationary throughout its exposure to cavitation. No mechanical acceleration stresses are applied to the specimen as there are when it is attached to a magnetostriction rod. As a specimen receives cavitation damage, small regions of reduced strength are produced and these small regions might be torn out of the specimen under mechanical acceleration. From the point of view of an analytic consideration of the cavitation bubble behavior, it also would seem advantageous to have a stationary boundary and no gross water velocity or acceleration.

For the photomicrographic observations and for the x-ray analyses of cavitation damage, the specimens were mounted flush in the base plate, Fig. 1. Under these conditions the cavitation cloud has a diameter of approximately 0.37 in., Fig. 2.

PHOTOMICROGRAPHIC OBSERVATIONS OF CAVITATION DAMAGE

A series of experiments were made to observe the optical changes in the surface of a specimen under cavitation damage. Fig. 6 shows the surface of a nickel specimen after a 5-min exposure to the cavitation cloud. It is of interest to note the unevenness of the damaged surface. This appearance of microscopic "hills and valleys" is characteristic of the behavior of very soft materials.



FIG. 4 BUBBLE CLOUD AT TWO DIFFERENT POINTS OF PRESSURE CYCLE

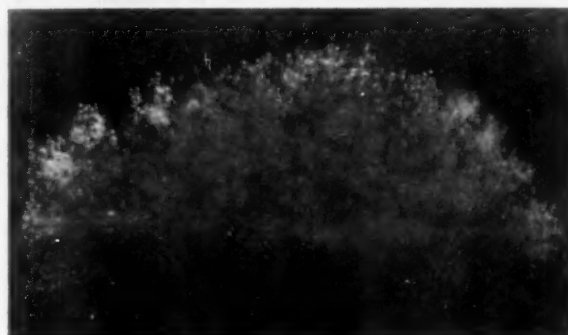
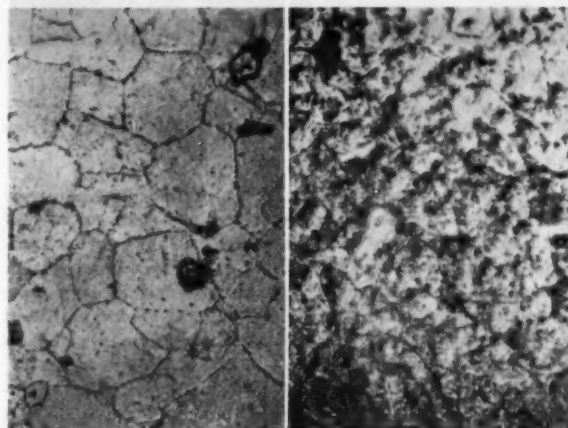


FIG. 5 BUBBLE CLOUD WITH TOP LIGHTING



(a) Before exposure to cavitation (b) After 5-min exposure to cavitation

FIG. 6 PHOTOMICROGRAPH OF NICKEL SPECIMEN; $\times 560$

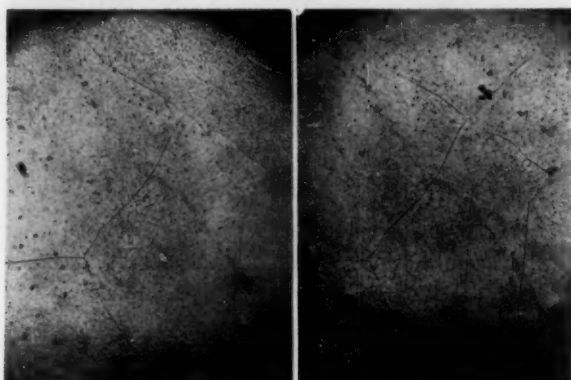
Fig. 7 shows a stainless-steel specimen which is of somewhat greater interest, metallurgically, since slip lines are clearly in evidence. These results may be compared with the photomicrographs of a harder material, an example of which is titanium 130-A, Fig. 8. No optical alteration in the surface of this material is evident even after prolonged exposure to the cavitation cloud.

This series of photomicrographs is typical of results with other materials. The very soft materials quickly show a microscopic hill-and-valley effect and not much can be deduced regarding



(a) Before exposure to cavitation (b) After 5-min exposure to cavitation

FIG. 7 PHOTOMICROGRAPH OF STAINLESS-STEEL SPECIMEN; $\times 560$



(a) Before exposure to cavitation (b) After 2-hr exposure to cavitation

FIG. 8 PHOTOMICROGRAPH OF TITANIUM SPECIMEN (WITH 7 PER CENT MANGANESE); $\times 200$

crystallographic changes. In materials of intermediate hardness, this hill-and-valley effect is not so evident and some evidence of crystallographic changes may be observed. In the very hard materials no change is evident. In these studies the cavitation damage was kept sufficiently light so that no significant weight loss in the specimens was obtained. Optical observations and photomicrography appear to be of limited value for a polycrystalline material which is the usual state of a specimen.

Photomicrography can be a powerful tool in revealing alterations in crystal structure if the specimen is a single crystal. A zinc monocrystal was made available to the authors through the kindness of Dr. Thad Vreeland, Jr., of the Engineering Division, California Institute of Technology. A basal plane (0001 plane) was obtained for exposure to cavitation by cleavage in liquid nitrogen. When this surface is stressed by the cavitation field, the hexagonal structure of the exposed plane becomes evident.

Fig. 9 shows this effect when the surface is exposed to a cloud of extent small compared with the exposed area. The nearly point application of the cavitation stress brings out the hexagonal structure very readily. Fig. 10 shows a basal plane in a zinc crystal before exposure to cavitation. Figs. 11, 12, and 13 show this same specimen after exposure to a broad cloud of cavitation covering the specimen. Similar observations have been made on a zinc crystal cut parallel to a twinning plane and the results are shown

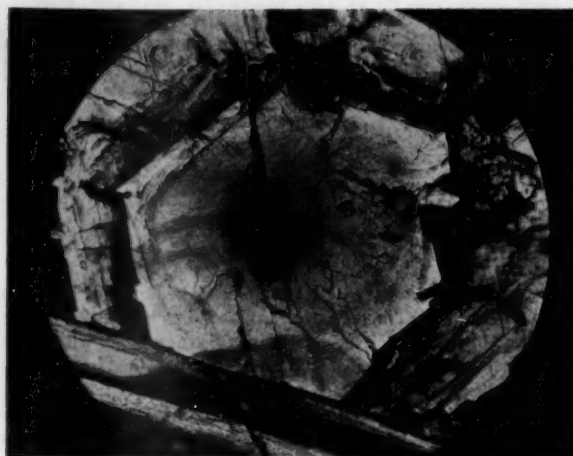


FIG. 9 CENTER OF DAMAGED AREA ON (0001) PLANE OF A SINGLE ZINC CRYSTAL AFTER 10-MIN EXPOSURE TO $1/16$ -IN-DIAM CAVITATION BUBBLE CLOUD; $\times 38$

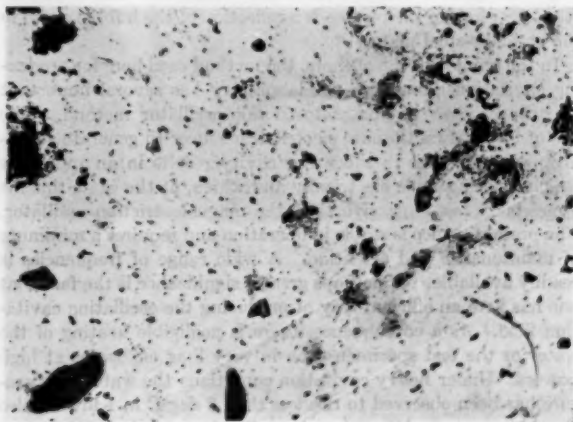


FIG. 10 ANNEALED SURFACE OF A ZINC MONOCRYSTAL CLEAVED ALONG A BASAL PLANE—UNDAMAGED; $\times 250$

in Figs. 14, 15, and 16. The geometry of the zinc crystal is sketched in Fig. 17.

While some of the previous photomicrographs gave evidence that cavitation damage is initiated through plastic deformation of the exposed specimen, very striking and detailed additional evidence of this mechanism are obtained from these studies on the monocrystal. In particular, severe plastic deformation of the zinc monocrystal is found in the form of twinning in the specimens exposed on a slip plane and in the form of slip in the specimens exposed on a twinning plane.

X-RAY OBSERVATIONS OF CAVITATION DAMAGE

Changes in structure can be observed readily in a polycrystalline specimen by the use of x-rays. For this reason a series of observations were taken on ordinary specimens exposed to cavitation damage. Fig. 18 shows x-ray diffraction patterns from a nickel specimen before exposure to cavitation, after 2-sec exposure to a broad cavitation cloud, and after 10-sec exposure. The blurring of the patterns, which is quite marked after only 2-sec exposure, shows that plastic deformation of the nickel microcrystals starts almost immediately upon exposure to cavitation. Fig. 19 shows similar results in brass. These specimens showed no significant

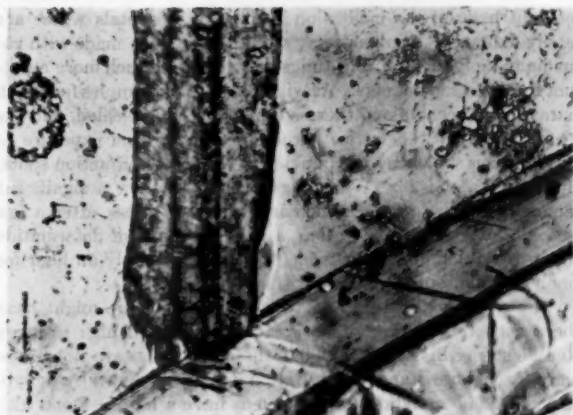


FIG. 11 ANNEALED SURFACE OF A ZINC MONOCRYSTAL CLEAVED ALONG A BASAL PLANE
(Note appearance of many new twin bands, cross twins, and hexagonal pits, 15-sec exposure; $\times 250$.)



FIG. 12 ANNEALED SURFACE OF A ZINC MONOCRYSTAL CLEAVED ALONG A BASAL PLANE
(Observe the increase in surface roughness and pitting, 30-sec exposure; $\times 250$.)



FIG. 13 ANNEALED SURFACE OF A ZINC MONOCRYSTAL CLEAVED ALONG A BASAL PLANE—DEVELOPMENT OF CAVITATION DAMAGE
(Note slip lines in twin bands and growth of pits 80-sec exposure; $\times 100$.)

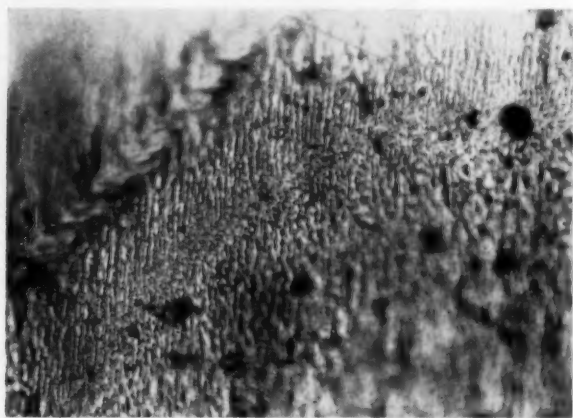


FIG. 14 POLISHED AND ANNEALED SURFACE OF A ZINC MONOCRYSTAL CUT PARALLEL TO A TWINNING PLANE
(Texture of surface is due to preferred attack of etchant. The ropy line across the surface is a twin band due to cutting the specimen, undamaged; $\times 500$.)

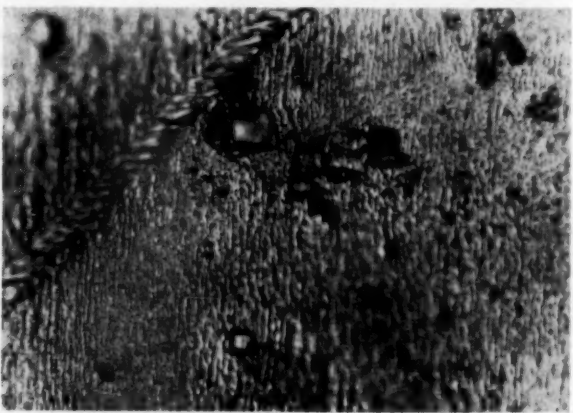


FIG. 15 POLISHED AND ANNEALED SURFACE OF A ZINC MONOCRYSTAL CUT PARALLEL TO A TWINNING PLANE
(Texture of surface is due to preferred attack of etchant. Ropy line across surface is a twin band due to cutting the specimen—note rectangular pits in the surface, 10-sec exposure; $\times 500$.)



FIG. 16 POLISHED AND ANNEALED SURFACE OF A ZINC MONOCRYSTAL CUT PARALLEL TO A TWINNING PLANE
(Texture of surface is due to preferred attack of etchant. Ropy line across surface is a twin band due to cutting the specimen, 210-sec exposure; $\times 500$.)

Geometry of a Zinc Crystal

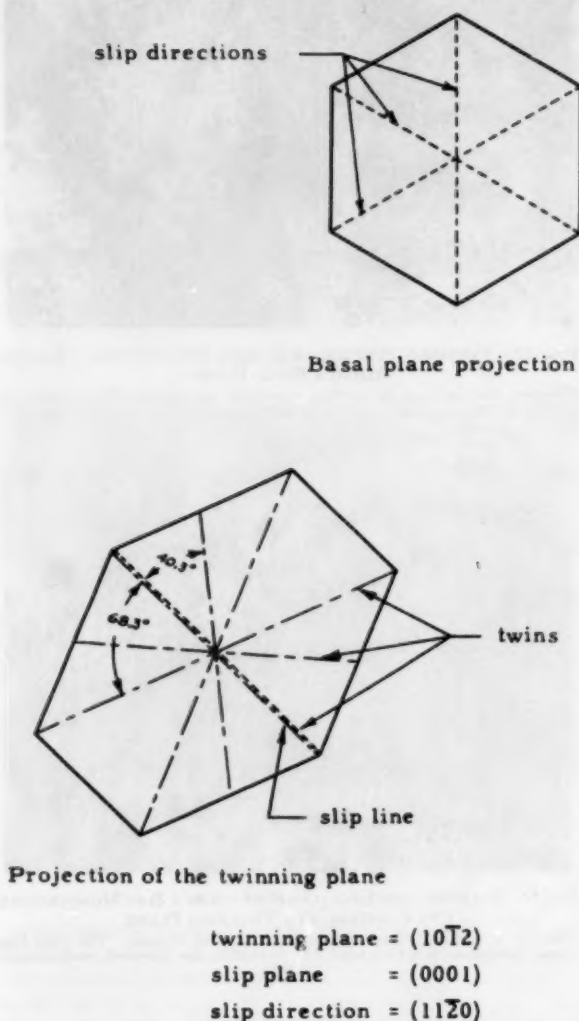


FIG. 17 GEOMETRY OF A ZINC CRYSTAL

optical change or weight loss; yet the alteration in structure is very evident from the x-ray observations.

To determine the depth of the region of plastic deformation, thin layers were removed from the surface until the x-ray pattern returned approximately to its original configuration. Layers were "peeled off" uniformly by the process of electrolytic polishing. These results are illustrated in Figs. 20 and 21 for nickel, and in Fig. 22 for brass. A similar study was made of a pure titanium specimen and is illustrated in Fig. 23. Here the x-ray diffraction spots are replaced by circular lines since the specimen was rotated while being exposed to the x-ray beam. The broadening of the circular lines shows again the plastic deformation which takes place on exposure to cavitation. To determine the depth of penetration of the region of plastic deformation in titanium, layers were removed in steps by etching until the lines returned approximately to their original sharpness.

The x-ray diffraction patterns in Figs. 18 to 23 show the effects

of light exposure to cavitation in a group of metals which are easily damaged. Similar x-ray studies have been made with vanadium, molybdenum, and tungsten which are much more resistant to cavitation damage. As with all the other samples for which x-rays have been taken, these specimens were annealed, and the deformation of the lattice structure arising from exposure to cavitation is shown by the blurring of the x-ray diffraction spots, Figs. 24, 25, and 26. While Figs. 20 and 22 show a significant amount of plastic deformation in nickel and brass after a few seconds' exposure to cavitation, a similar amount of deformation takes several minutes for vanadium, approximately an hour for molybdenum, and several hours for tungsten.

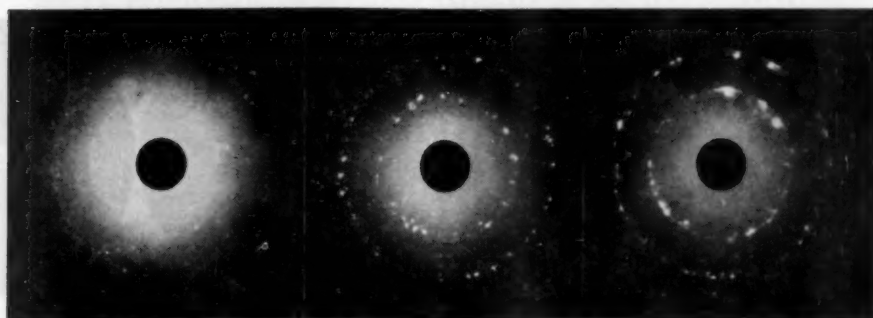
The resistance of tungsten to cavitation damage might have been expected in view of its great hardness and high ultimate tensile strength, Table 2. That these properties are not necessary for high resistance to cavitation damage is shown by molybdenum. The molybdenum specimens have a hardness and ultimate tensile strength similar to those of brass, and yet were very much more resistant to cavitation damage. The x-ray studies presented here show that cavitation damage takes place through plastic deformation, or cold-work, of the exposed solid. In some materials, such as nickel or brass, this plastic deformation occurs very quickly; in others, such as molybdenum or tungsten, relatively long exposures are required. The mechanism of plastic deformation demonstrates that the essential property for resistance to cavitation damage is the fatigue resistance of the solid. Further, the fatigue behavior of a solid is related in a complex way to its usually measured properties. It also must be kept in mind that the stresses applied to the solid by cavitation are of very short duration so that the fatigue behavior under low frequency stress application is not necessarily a measure of resistance to cavitation damage. The exceptional resistance of molybdenum to cavitation damage in view of its hardness and tensile strength may be understood from the known lag in the yielding of this solid to an applied stress. Thus, for the short stress pulses produced in cavitation, the deformation of molybdenum may be expected to be reduced.

RATE OF MATERIAL LOSS IN CAVITATION DAMAGE

A series of experiments were undertaken to secure a rough measure of resistance to cavitation damage in a group of polycrystalline substances. This series has a definite engineering interest. To get significant damage quickly, the specimens were made with a small tip protruding above the level of the bottom flat plate. By this means the cavitation cloud was concentrated over the end of the tip and did not appear elsewhere; the diameter of the region over which the cavitation cloud extended under these conditions was approximately 0.07 in. With this configuration, the cavitation cloud drilled nearly cylindrical holes in the specimens. The depth of these cylindrical holes for given exposure time is a measure of the resistance to damage. Measurements of this kind are summarized in Table 2. A photograph of a stellite specimen exposed to this concentrated cloud is shown in Fig. 27, and a corresponding photograph of a titanium 150-A specimen is shown in Fig. 28.

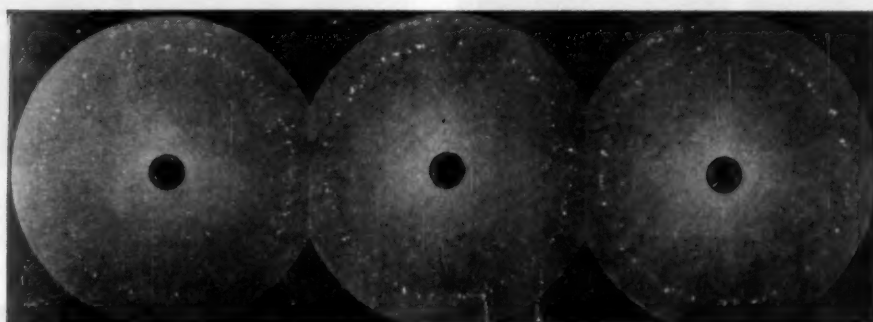
CHEMICAL EFFECTS IN CAVITATION DAMAGE

There has been considerable speculation that cavitation damage is primarily the result of chemical action. Such a corrosion picture of the damage is in contradiction with the view presented here of the damage arising from cold-work. In order to give further information on this question, a series of experiments were performed in which the cavitation cloud was generated in liquid toluene from which any dissolved air was completely removed. At the same time, helium at 1 atm was maintained above the liquid. Toluene is known to be a very inert liquid chemically and



(a) Before exposure (b) After 2-sec exposure (c) After 10-sec exposure

FIG. 18 X-RAY DIFFRACTION PATTERN OF NICKEL SPECIMEN SHOWING RAPID ONSET OF COLD-WORK ON EXPOSURE TO CAVITATION IN WATER



(a) Before exposure (b) After 2-sec exposure (c) After 10-sec exposure

FIG. 19 X-RAY DIFFRACTION PATTERN OF BRASS SPECIMEN SHOWING RAPID ONSET OF COLD-WORK ON EXPOSURE TO CAVITATION

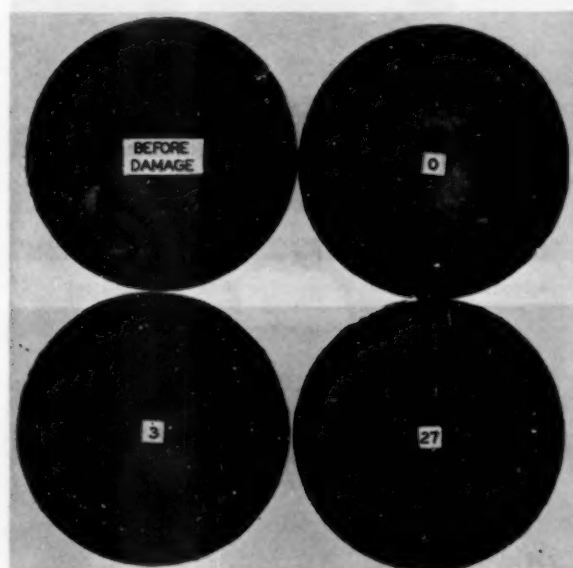


FIG. 20 ANNEALED NICKEL—DAMAGED 30 SEC—DEPTH OF LAYER REMOVED AFTER DAMAGE IS SHOWN IN MICRONS

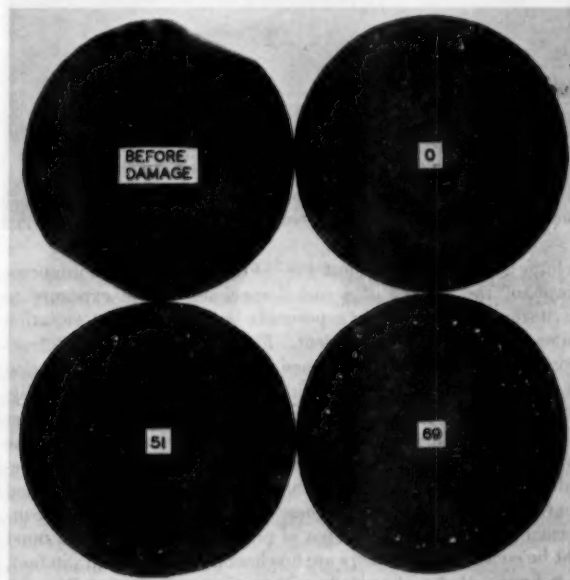


FIG. 21 ANNEALED NICKEL—DAMAGED 30 MIN—DEPTH OF LAYER REMOVED AFTER DAMAGE IS SHOWN IN MICRONS

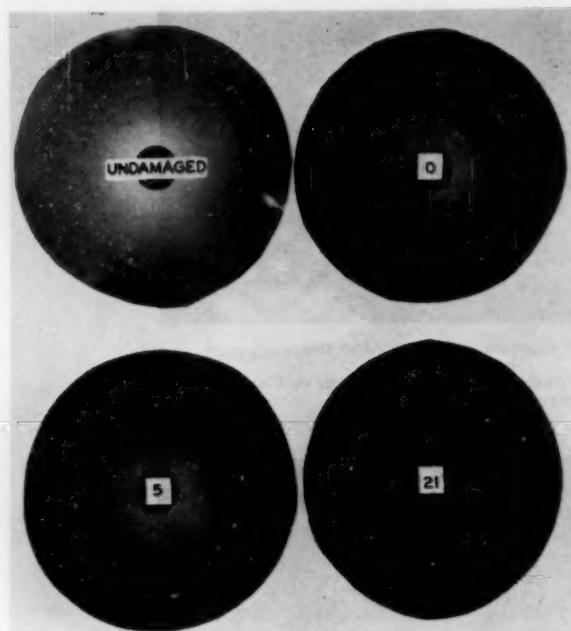


FIG. 22 ANNEALED BRASS—DAMAGED 30 MIN—DEPTH OF LAYER REMOVED AFTER DAMAGE IS SHOWN IN MICRONS



FIG. 23 PURE ANNEALED TITANIUM—DAMAGED 30 MIN—DEPTH OF LAYER REMOVED AFTER DAMAGE IS SHOWN IN MICRONS

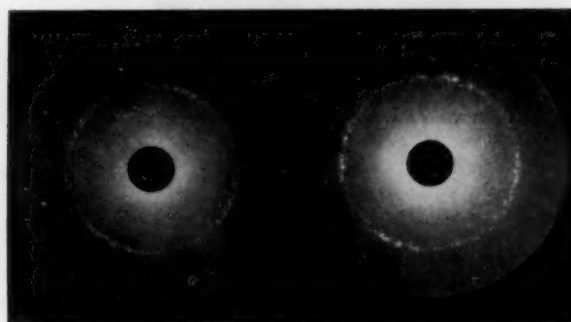


FIG. 24 X-RAY DIFFRACTION PATTERN OF VANADIUM SPECIMEN SHOWING COLD-WORK ON EXPOSURE TO CAVITATION IN WATER

helium is, of course, an inert gas. Fig. 29 shows a photomicrograph of the surface of a nickel specimen before exposure to cavitation and after 5-min exposure in liquid toluene. Cavitation damage to the surface is evident. Fig. 30 shows an x-ray diffraction picture of a nickel specimen before exposure to cavitation and after 2-sec exposure in liquid toluene. A comparison of Fig. 30 and the corresponding x-ray pattern obtained with nickel after 2-sec exposure in water, Fig. 18, shows that essentially the same type and extent of plastic deformation take place. Similar results have been found with stainless steel. It is to be concluded that, in so far as the basic process of cavitation damage is concerned, chemical effects are not of primary significance. It must not be concluded that there are no chemical effects at all; in fact, it is known that a chemically active environment can affect the fatigue properties of a solid material.

SUMMARY AND CONCLUSION

The apparatus for this study of cavitation damage, in the authors' opinion, is an inexpensive and efficient device for cavi-

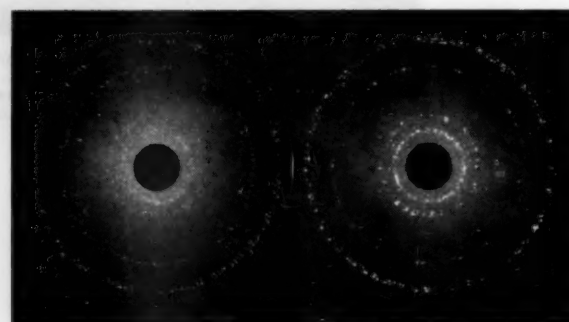


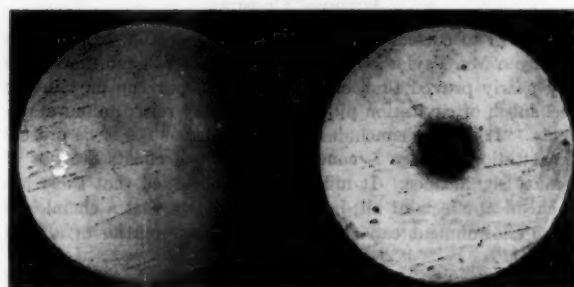
FIG. 25 X-RAY DIFFRACTION PATTERN OF MOLYBDENUM SPECIMEN SHOWING COLD-WORK ON EXPOSURE TO CAVITATION IN WATER



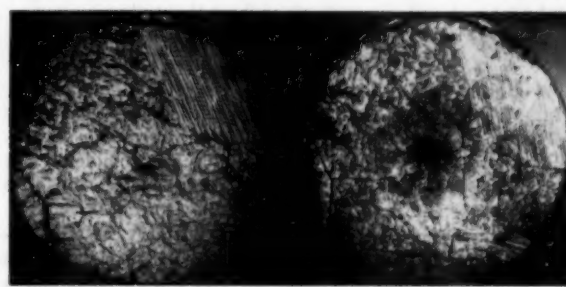
FIG. 26 X-RAY DIFFRACTION PATTERN OF TUNGSTEN SPECIMEN SHOWING COLD-WORK ON EXPOSURE TO CAVITATION IN WATER

TABLE 2 CHARACTERISTICS OF VARIOUS MATERIALS AND SUMMARY OF CAVITATION DAMAGE SUSTAINED IN TESTS

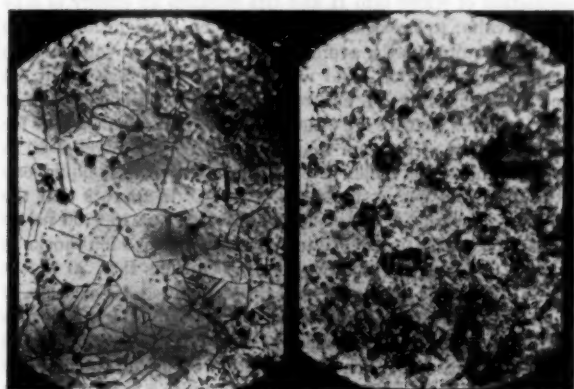
Material (arranged in order of increasing hardness) composition, per cent	Hardness, Bhn	Ultimate tensile strength (psi $\times 10^3$)	Modulus of elasticity (psi $\times 10^6$)	Depth of cavitation damage hole in microns (10^{-4} cm)						
				10 sec	1 min	15 min	30 min	1 hr	2 hr	3 hr
Aluminum (soft).....	16	16	10	10	80					
Titanium (annealed).....	58	79	16			43	78			
Nickel.....	90	50	30			80	115			
Molybdenum.....	120	57	50				10	25	60	100
Brass, Cu 70, Zn 30.....	123	56	13		1	85	128			
Stainless steel, Cr 18, Ni 8.....	163	102	29			15	28			
Titanium 75-A.....	203	80	16					30	66	
Steel (4130).....	258	130	30					32	55	
Tungsten.....	350	597	51				0	0	3	12
Titanium 130-A, Ti 92, Mn 7.9.....	351	130	16				0	3	16	
Colmonoy.....	400	61					0	3	18	34
Titanium 150-A, Ti 90, Cr 2.7, Fe 1.3.....	437	150	16				0	0	3	26
Stellite, Co 55, Cr 33, W 6.....	495	100	36				0	3	14	29
Pyrex.....	Moh 5		10			120				
Fused quartz.....	Moh 7		9			100				



(a) Undamaged (b) Damaged 1 hr

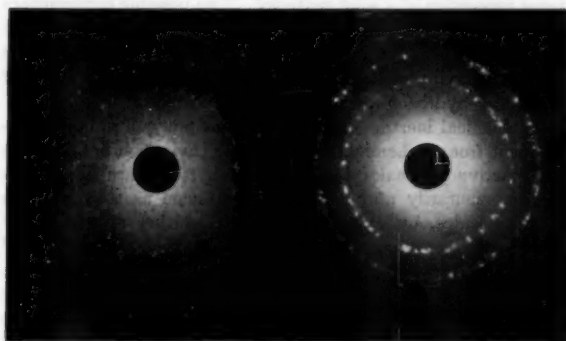
FIG. 27 STELLITE SPECIMEN—MAGNIFICATION $\times 40$ 

(a) Undamaged (b) Damaged 1 hr

FIG. 28 TITANIUM 150-A SPECIMEN—MAGNIFICATION $\times 40$ 

(a) Before exposure (b) After 5-min exposure

FIG. 29 ANNEALED AND ETCHED SURFACE OF NICKEL SPECIMEN EXPOSED TO CAVITATION IN LIQUID TOLUENE IN A HELIUM ATMOSPHERE



(a) Before exposure (b) After 2-sec exposure

FIG. 30 X-RAY DIFFRACTION PATTERN OF NICKEL SPECIMEN SHOWING RAPID ONSET OF COLD-WORK ON EXPOSURE TO CAVITATION IN TOLUENE IN A HELIUM ATMOSPHERE

tion research or testing. Cavitation tests are extremely simple, and relative resistance to cavitation damage such as is given in Table 2 is very easily obtained.

Ordinary polycrystalline specimens and pure monocrystals have been exposed to cavitation damage. Photomicrographic examination and x-ray diffraction patterns both show that plastic deformation occurs. This plastic deformation appears to set in almost immediately in soft substances such as nickel, brass, or pure titanium. Since these are substances with ultimate tensile strengths of the order of 50,000 psi, it is indicated that cavitation stresses may be at least of this magnitude. On the other hand, the very slow onset of damage in a material such as tungsten or

titanium 150-A with ultimate tensile strengths of the order of 130,000 psi or greater might be taken as an indication that cavitation stresses lie below this magnitude.

It should be emphasized that cavitation damage represents the interaction of a complex of properties of the solid so that consideration of ultimate tensile strengths alone, or of hardness alone, or any other single property would be misleading. The general mechanism of the damage does appear to be reasonably clear. A polycrystalline or monocrystalline specimen exposed to cavitation undergoes eventual or almost immediate plastic deformation, depending on its hardness and yield properties. This plastic deformation is a cold-work of the material which leads to fatigue and failure of portions of the specimen.

The similarity in the plastic deformation produced in a variety

of specimens with different chemical properties makes it unlikely that chemical action is of particular significance in cavitation damage, at least as produced here. Conclusive additional evidence on this point is furnished by the photomicrographic and x-ray studies of cavitation damage produced in air-free toluene with a helium atmosphere. Toluene and helium are particularly inert substances chemically. These experiments show the same damage effects as are found in water.

ACKNOWLEDGMENTS

This study was supported by the U. S. Navy Office of Naval Research, Mechanics Branch. The authors wish to express their appreciation to Mr. J. E. Neimark for valuable assistance with the observations on zinc monocrystals. The authors are also greatly indebted to Profs. Pol Duwez and David Wood of the Engineering Division of the California Institute of Technology, not only for making facilities available to them but also for their helpful interest and comments throughout the study.

Discussion

IRVING TAYLOR.* The authors of this splendid paper have very nearly proved that chemical actions play no part in the mechanism of cavitation pitting. Before taking this completely for granted, it is advisable to consider two plausible phenomena that could be important, if secondary.

Not only the most resistant materials but also the medium-resistant materials (vanadium, molybdenum, 4130 steel, stainless steel) showed "incubation periods," in Table 2 of the paper and in the x-ray studies. If these delays are required for fatigue, then what is the mechanism of the fatiguing? The authors have allowed that a chemically active environment can affect the fatigue properties of a solid material. Two possible contributors that may need to be considered are (a) partial dissociations of the liquid to free radicals and chemical attack on portions of the solid, (b) partial dissociations to free radicals and penetration of the solid by hydrogen atoms.

The chemical inertness of a liquid does not guarantee an environment free of chemical activity once cavitation occurs. Intensely active molecular fragments of short life are likely produced momentarily and locally from toluene as well as from water molecules, by partial dissociations, either at the cavity collapse-points or along the solid surface by shock-wave impacts.[†] These fragments include free atoms and free radicals such as H·, ·OH, ·OOH, ·CH₃, ·C₂H₅, ·C₄H₉CH₃, note that these are not ions. Free radicals are known to play vital roles in extremely fast chemical reactions such as flames and explosions, and are reported to cause adjacent stable substances to break down or become labile along with them. A stable metal surface (particularly of a good catalyst like nickel, iron, copper, or palladium) probably becomes so activated in the local presence of radicals from cavitation that its atoms participate in chemical reactions impossible elsewhere. Other reactions would concurrently restore the metal atoms, but not necessarily in the same exact place.

Also, there is a likelihood of penetration into the solid, by some of the free hydrogen atoms or by hydrogen molecules. Inter-

stitial hydrogen seems to make titanium alloys become brittle. This possibility of hydrogen penetration—through impacts and shock waves, partial dissociations, more impacts and shock waves—could well be an intermediate step between the beginning of cavitation and the finding of cold-work or deformation. The hydrogen inside the lattice, some of it still active chemically, could bring about "unbindings" or "loosening" of metal atoms one from another. Would not the x-ray patterns also be changed before any microscopic cracks could be observed?

Despite the weight of the authors' evidence there remains some uncertainty that the first deformations after exposure to cavitation are due to cavitation stresses alone—at least on the more resistant materials. As in the somewhat analogous phenomenon of fretting corrosion analyzed by H. H. Uhlig,[‡] a combination of mechanical and chemical factors may comprise the mechanism of cavitation damage.

AUTHORS' CLOSURE

We wish to correct a possible misinterpretation which might arise from Mr. Taylor's opening remark that "The authors have very nearly proved that chemical actions play no part in the mechanism of cavitation pitting." We may reiterate our statement: "It is to be concluded that, in so far as the basic process of cavitation damage is concerned, chemical effects are not of primary significance. It must not be concluded that there are no chemical effects at all; in fact, it is known that a chemically active environment can affect the fatigue properties of a solid material."

The point under discussion is whether distilled water or liquid toluene are chemically active environments or not. Mr. Taylor proposes that these inert liquids become active through dissociation into free atoms or radicals. It is to be observed that the dissociation of water into H and OH requires an energy of 5.14 electron volts; to break a carbon-hydrogen bond in an organic molecule requires on the average about 4.3 electron volts, and to break a carbon-carbon bond requires on the average about 3.4 electron volts. Since 1 electron volt = 23.6 kilocalories per mole, it follows that we are concerned with activation energies for the dissociation of water or toluene in the neighborhood of 80 to 100 kilocalories per mole. Now a shock wave in a liquid can only supply this energy by shock heating. It appears that the shocks of concern have pressure jumps of about 100,000 psi or less and a computation of the temperature rise for such shocks in water or toluene gives a temperature rise of less than 20 C. It follows that these shock waves cannot produce dissociation of the liquid. The possibility remains that higher temperatures may be encountered in the vapor of the cavitation bubbles when they are collapsed. Even if dissociation occurs at the point of collapse, the lifetime of the dissociation products is so small that it may be presumed that recombination would take place before they could diffuse through the liquid to the solid surface.

Our observations all point to the interpretation of cavitation damage as arising primarily from repeated application of high stresses of short duration which accompany vapor bubble collapse. Any attempt to explain the phenomenon by chemical dissociation in the liquid would seem to be based on unlikely speculations.

*Senior Mechanical Equipment Engineer, The Lummus Company, New York, N. Y. Mem. ASME.

†"Cavitation Pitting by Instantaneous Chemical Action From Impacts," by Irving Taylor, ASME Paper No. 54-A-109.

‡"Mechanism of Fretting Corrosion," by H. H. Uhlig, *Journal of Applied Mechanics*, Trans. ASME, vol. 76, 1954, pp. 401-407.

Secondary Flow in Axial-Flow Turbomachinery¹

By L. H. SMITH, JR.,² CINCINNATI, OHIO

A method is presented by which one can include the principal effects of secondary flow in the design of multi-stage axial-flow turbomachinery. The method uses the conventional axial-symmetry solution as a basic flow picture from which the induced turning and radial velocities characteristic to the secondary flow can be deduced. Frictional effects, such as the growth of boundary layers on the blades, are excluded from the analysis, but the disposition of the fluid which comes from boundary layers on the hub and casing surfaces upstream of a blade row can be accounted for in an approximate manner by assuming that the fluid is frictionless while in the blading. To get a first-order approximation of the secondary flow in a blade row, it is necessary to know only the number of blades in the row and the axisymmetric-flow solution. The principal assumption involved in this approximation is that the distortion of the flow surfaces due to the secondary velocities is small. The method is sufficiently accurate for predicting the secondary flows which occur as a result of nonuniform blade loading, but solutions for boundary-layer flows must be considered only qualitative because of the large flow-surface distortions which are associated with these flows.

NOMENCLATURE

The following nomenclature is used in the paper:

- a = distance between dividing streamlines (Fig. 3), ft
- b = distance between hub and tip measured along normal direction, ft
- c = distance along upstream streamline (Fig. 3), ft
- f = distance related to time lag between particles (Fig. 1), ft
- k_i = coefficient in secondary vorticity expansion, Equation [25], 1/sec
- m = constant controlling boundary layer secondary vorticity, Equation [27], dimensionless
- n = distance normal to axisymmetric stream surfaces, ft
- N = number of blades in a row, dimensionless
- r = radial co-ordinate, ft
- S = strength of source field, 1/sec
- t = time, sec
- Δt = time difference for particles traversing an airfoil, Equation [1], sec
- \vec{V} = absolute fluid velocity, fps

- \vec{V}' = velocity induced by source field, fps
- \vec{V}_s = secondary velocity causing vortex sheet, fps
- \vec{W} = fluid velocity relative to blade system being considered, fps
- x = distance normal to streamlines and normal to n -direction, ft
- y = distance in n -direction, ft
- α = a/b , dimensionless
- β = angle between relative velocity and meridional stream line, dimensionless
- Γ = circulation around moving circuit, Equation [3], ft²/sec
- Γ_v = vane circulation, $\Gamma_v = \frac{2\pi}{N} (r_1 V_{\theta 1} - r_2 V_{\theta 2})$; for applications, Γ_v is always considered to be positive, ft²/sec
- $\vec{\zeta}$ = absolute vorticity of axisymmetric flow, $\vec{\zeta} = \nabla \times \vec{V}$, 1/sec
- $\vec{\zeta}_s$ = secondary vorticity; i.e., vorticity contributing to cellular pattern of secondary circulations, 1/sec
- η = y/b , dimensionless
- η_1 = y/a , dimensionless
- λ = angle of distortion of a stream surface, dimensionless
- ξ = x/a , dimensionless
- ρ = mass density, lb-sec²/ft⁴
- σ = distance along a streamline, ft
- ϕ = velocity potential for source field flow, ft²/sec
- ψ = stream function for secondary flow, ft²/sec

Subscripts:

- \perp = component perpendicular to relative velocity
- ∞ = vector mean of upstream and downstream conditions
- A = actual flow (primary + secondary)
- b = blade-to-blade flow
- m = meridional component
- x = component of secondary quantity in x -direction
- y = component of secondary quantity in y -direction
- θ = tangential component
- ξ = component of dimensionless secondary quantity in x -direction
- σ = component of secondary quantity in primary streamline direction
- 1, 2 = stations before and after a blade row

INTRODUCTION

The flow of a frictionless adiabatic fluid through a turbomachine is described by the Euler equations of motion, which are special cases of the Navier-Stokes equations, for the viscosity equal to zero. If the fluid entering the machine is everywhere irrotational, it will remain irrotational regardless of the shapes or motions of the passages through which it moves, except on a number of surfaces called vortex sheets where a discontinuity in velocity exists. These vortex sheets will be shed from the trailing edges of the moving and stationary vanes when the vane circulation is not uniform and/or when the relative flow approaching the vanes is not steady. If there are at least two blade rows in the machine, one

¹ The work on which this paper is based was done at the Johns Hopkins University, Baltimore, Md., under the sponsorship of Wright Air Development Center, Dayton, Ohio. The author wishes to acknowledge the stimulation and guidance given him by Drs. George F. Wislicenus and Hsuan Yeh during the course of the work.

² Compressor Aerodynamics Engineer, Aviation Gas Turbine Division, General Electric Company. Assoc. Mem. ASME. Contributed by the Compressor Committee of the Hydraulic Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 7, 1954. Paper No. 54-A-158.

of which is moving relative to the other(s), the flow will never be steady and hence there will always be vortex sheets created. The picture is therefore quite complex even in the simplest cases.

The far-reaching assumption which is generally made in order to permit a reasonably simple analysis is the assumption that each blade row in the machine is composed of an infinite number of infinitely thin blades. This reduces the problem to one of two independent variables rather than four, since all physical quantities become uniform in the tangential direction and the flow becomes steady. This is called the assumption of axial symmetry. In the axisymmetric picture we are able to visualize axisymmetric stream surfaces which are obtained by revolving a fluid streamline about the axis of rotation.

It is usually assumed, when designing the finite number of blades to give the desired flow distribution, that the fluid remains on the axisymmetric stream surfaces. On this basis two-dimensional blade-to-blade solution procedures have been developed; for instance, those given by Stanitz and Ellis (1)³ and Wu, Brown, and Prian (2). These methods give the velocity fields around the blade profiles under the assumption that there is no flow normal to the axisymmetric stream surfaces which were obtained from the infinite blade solution. For axial-flow machines, theoretical and experimental two-dimensional (plane) cascade methods may be used as an approximation for the blade-to-blade flow.

It is pertinent to inquire about the discrepancies between the actual three-dimensional flow and the flow which is obtained when one uses the two-dimensional procedures just described. One discrepancy is caused by friction in the fluid; the boundary layers which build up on the blades certainly do not act as one would predict from two-dimensional theories. A second occurs because of the unsteadiness of the flow which is caused by the relative motion of adjacent blade rows. Still a third discrepancy occurs because, in order to satisfy the equilibrium equations, the flow does not actually remain on the axisymmetric stream surfaces. The object of this paper is to provide a means for taking into account quantitatively this third discrepancy. Thus, it is desired to determine the velocities of the fluid particles normal to the axisymmetrically determined stream surfaces and the tangential and meridional velocities which are associated with them.

In order to isolate the effect which is to be investigated, it is assumed that the flow is frictionless while passing through any given blade row, and that the flow is axisymmetric as it enters each row. Let the term "primary flow" be given to the flow which one obtains by using successively the axisymmetric and blade-to-blade methods, and let the term "secondary flow" denote the difference between the actual flow and the primary flow.

Although friction has been excluded from the analysis in each individual blade row, it is not necessary that the cumulative effects of friction be neglected. These may indeed be included in the initial axisymmetric solution, and if this is done realistically, the typical hub and casing boundary layers of the latter stages will appear. The secondary flow as defined will, therefore, include the boundary-layer secondary flow, which has received much attention from many writers lately, as well as the secondary flow which results from nonuniform blade loading.

Method for Finding Secondary Velocities. In order to find the secondary flow, we must be able to solve the problem of the flow with vorticity in a curved channel. The channel is the space bounded by two blades and the hub and casing surfaces. This problem has received a fair amount of attention, for instance from Squire and Winter (3), Hawthorne (4), Kronauer (5), Eichenberger (6), and Ehrich (7). These investigators obtained

solutions by solving (more or less approximately) for the flow field inside the channel, using the channel shape and the upstream flow as boundary conditions. For general channel shapes this type of procedure is obviously quite difficult.

By using the fact that the channel is bounded on two sides by airfoils, and using the circulation around these airfoils as part of the boundary conditions, we can arrive at a more tractable method of solution. This adds no new complication since the vane circulation Γ_v is known (excluding secondary-flow effects) once the tangential velocity distributions and the number of vanes are specified (see Γ_v in Nomenclature).

Before going into the details of the method, it will be well to consider the physical nature of the problem. A vane turns the fluid because there is a pressure difference across the thickness of the vane. Now if this pressure jump is different for different sections along the span of the vane, there will be a pressure gradient along each face of the vane in the spanwise direction, in addition to the pressure gradient caused by axisymmetric equilibrium. This gradient causes spanwise velocities which contribute to the secondary flow.

In order to have constant loading along the span (and thus have no induced spanwise velocities), we must have the product of the vane circulation times the oncoming meridional velocity times the fluid density ($\Gamma_v V_{m\rho}$) constant along the span, since this product gives the tangential blade force and thereby the pressure jump across the vane. This statement is of course only qualitatively true since the pressure distribution for a given force is dependent on the axial blade length and loading distributions.

The spanwise (normal) velocities (which have different directions on the pressure and suction sides of the vane) create a trailing vortex sheet. The strength of this sheet is proportional (approximately) to

$$\frac{1}{V_{m\rho}} \frac{d}{dn} (\Gamma_v V_{m\rho})$$

and not proportional to $d\Gamma_v/dn$ as has been sometimes implied.

Although an insight into the nature of secondary velocities is obtained most easily by a consideration of pressure gradients, the actual calculation of these velocities is more conveniently done by use of the vortex laws of fluid mechanics. We find, however, that these laws usually require an a priori knowledge of the fluid motion. For instance, one of the Helmholtz laws states that a vortex line moves with the fluid. Kelvin's law says that the circulation around a circuit which moves with the fluid remains constant. It is therefore necessary to know ahead of time, at least approximately, the primary fluid motions in order to calculate the secondary motions.

This brings us to the following method of approximation:

We assume initially that the flow behaves as if it were to remain on the axisymmetric stream surfaces. We solve for the primary flow picture by using blade-to-blade methods, or, if we have an axial-flow machine, by using two-dimensional cascade data corrected for changing axial velocity. With this flow as a basis, we invoke the laws of vortex motions to calculate secondary velocities.

The fluid is assumed to be frictionless while going through the system, regardless of how its vorticity came to exist originally.

Instead of solving for the complete primary flow picture as mentioned in the foregoing method, we need only deduce, for practical purposes, one particular characteristic of the flow. This characteristic is contained in the following problem:

Consider the two-dimensional problem of the irrotational flow of a perfect fluid around an airfoil, see Fig. 1. We ask the question: Given the particle *A*, where is the particle *B* which will be at the trailing edge of the vane when *A* gets there? Or we might ask

³ Numbers in parentheses refer to the Bibliography at the end of the paper.

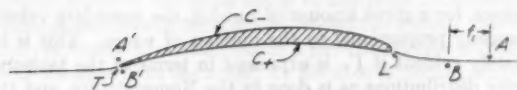


FIG. 1 LOCATION OF PARTICLES TRAVERSING AN AIRFOIL

equivalently: How much more time does it take for a particle to get from the leading edge to the trailing edge of an airfoil if it goes under rather than over it?

Let V = velocity

Δt = time difference (bottom—top)

σ = distance along airfoil

Then

$$\Delta t = \int_{C_+}^T \frac{d\sigma}{V} - \int_{C_-}^T \frac{d\sigma}{V} \dots \dots \dots [1]$$

This problem has been considered in some detail by the author in references (8) and (9). The matter is complicated by the fact that particles which approach the airfoil along its stagnation streamline require an infinite amount of time to reach the blunt leading edge. By restricting applications to thin airfoils, and/or by choosing particles which are not exactly on the stagnation streamline, it was found that the formula

$$\Delta t = \frac{\Gamma_v}{V_\infty^2} \dots \dots \dots [2]$$

represents a good approximation for conventionally shaped airfoils either isolated or in cascade.

The manner in which this information about the primary flow is used will be discussed in the following sections. There are two avenues of approach which can be used to calculate secondary velocities. The first, which we shall call the "airfoil theory," treats the blades as airfoils; the secondary flow is induced by the vortex sheets which are attached to the vane trailing edges. The other method which we shall call the "channel theory" has been described briefly in the foregoing in connection with several workers who have used it. We shall discuss the airfoil theory first.

THE AIRFOIL THEORY

In this section a method for calculating the characteristics of the vortex sheets which are shed from the trailing edges of the blades will be presented. The well-known Prandtl wing theorem which states that the strength of the vortex sheet shed by a wing is equal to the spanwise gradient in circulation along the wing will not generally be valid for turbomachinery, because this theorem assumes that the flow approaching the wing is irrotational. We will therefore investigate how this theorem must be modified so that it will be applicable to rotational flows. For the present purposes we will be interested in the case when the flow upstream of the vane is axially symmetric. Thus the flow will be steady relative to the vane and the vorticity vector upstream will lie in a plane which is normal to the span of the vane.

Formulation of the Theory. We focus attention on one of the airfoils in a cascade, moving or stationary, Fig. 2. The oncoming fluid (frictionless but compressible) has distributed vorticity but no entropy gradients. Consider the circuit $ABCDEF$ in the oncoming fluid which is chosen so that at some time later it will wrap itself around the vane in the position $A'B'C'D'E'F'A'$. The lines $A'B'$, $D'C'$, $A'F'$, and $D'E'$ are selected to lie in the axisymmetric stream surfaces of the primary flow solution. If the vane has a blunt nose the particles A and D take an infinite

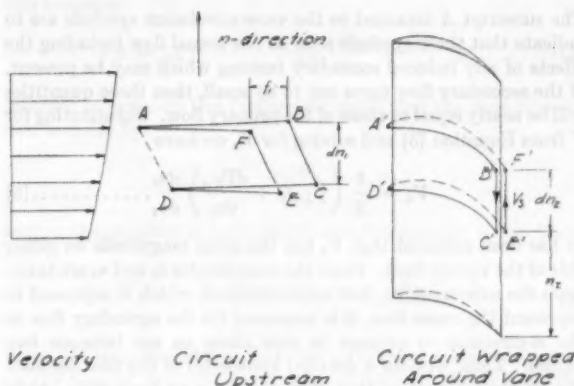


FIG. 2 MOVING CIRCUIT UPSTREAM OF VANE AND, AT LATER TIME, ON VANE

amount of time to get to the vane. For this kind of vane it is necessary to think of the circuit as always being an infinitesimal distance from the vane surface. However, if the profile nose is wedge-shaped or cusp-shaped or, if the whole profile is infinitely thin with the fluid flowing smoothly across the leading edge, then we may take a circuit which touches the vane. The line AF does not coincide in direction with the line AB in Fig. 2 because it is desired to have B' next to F' at the trailing edge of the vane. Therefore, due to the spanwise secondary motion, particles F and B must come from different streamlines.

Kelvin's theorem says that the absolute circulation around this circuit will remain constant.⁴ In general there will be a circulation around the circuit because the upstream flow contains absolute vorticity. This circulation can be calculated by integrating the vorticity over any area bounded by the circuit according to Stokes' Theorem. The circuit upstream, Fig. 2, is folded over on itself along the line AD so that the contribution of the area $AFED$ to the circulation integral is just canceled by the overlapping part of the $ABCD$ area. The fact that F and B are not on the same streamline is neglected here, but this results in no appreciable error since E and C are displaced by nearly the same amount. We integrate over the remaining area $FBCE$ to get the circulation around the circuit. This gives the result

$$\Gamma = \int_{\perp} \zeta_{\perp} f dn_1 \dots \dots \dots [3]$$

in which ζ_{\perp} is the component of the absolute vorticity which is perpendicular to the velocity. We shall consider Γ and ζ_{\perp} positive when they tend to cause the secondary velocity V_s to be in the n -direction on the suction side of the vane, as shown in Fig. 2. This will happen whenever the vector represented by the vector product $\vec{W} \times \vec{\zeta}$ points in the n -direction. The vane circulation will always be taken to be positive.

By stating that the whole is equal to the sum of its parts we have, for the circuit wrapped around the vane

$$\Gamma = \Gamma_{C'D'E'} - \Gamma_{B'A'F'} + 2V_s dn_2$$

Noting that the first two of these are actually vane circulations, we can write

$$\begin{aligned} \Gamma &= \Gamma_{VA}(n_2) - \left[\Gamma_{VA}(n_2) + \frac{d\Gamma_{VA}}{dn_2} dn_2 \right] + 2V_s dn_2 \\ \Gamma &= - \frac{d\Gamma_{VA}}{dn_2} dn_2 + 2V_s dn_2 \dots \dots \dots [4] \end{aligned}$$

⁴ See, for instance, Vassonyi (10).

The subscript A attached to the vane-circulation symbols are to indicate that these symbols refer to the actual flow including the effects of any induced secondary turning which may be present. If the secondary flow turns out to be small, then these quantities will be nearly equal to those of the primary flow. Substituting for Γ from Equation [3] and solving for V_s we have

$$V_s = \frac{1}{2} \left(\zeta_{\perp} f_1 + \frac{d\Gamma_{VA}}{dn_1} \frac{dn_1}{dn_2} \dots \dots \dots [5] \right)$$

It has been assumed that V_s has the same magnitude on either side of the vortex sheet. Since the co-ordinates n_1 and n_2 are taken from the axisymmetric flow approximation, which is supposed to represent the mean flow, it is necessary for the secondary flow in the n -direction to average to zero along an arc between two blades. Thus, without a detailed knowledge of the flow between the blades, we assume that V_s is the same on both sides of the sheet.

The distance f_1 we know to be

$$f_1 = W_1 \Delta t = W_1 \frac{\Gamma_{VA}}{W_{\infty}^2} \dots \dots \dots [6]$$

in which W is the velocity relative to the vane and Δt is the time difference for the two particles to go around the vane as given by Equation [2]. Substituting f_1 from Equation [6] into [5] we obtain

$$V_s = \frac{1}{2} \left(W_1 \zeta_{\perp} \frac{\Gamma_{VA}}{W_{\infty}^2} + \frac{d\Gamma_{VA}}{dn_1} \frac{dn_1}{dn_2} \right) \dots \dots \dots [7]$$

This equation gives the secondary velocity which creates the trailing vortex sheet. It is seen to be made up of two parts, one being proportional to the spanwise gradient in vane circulation as in the Prandtl wing theory, and the other being a function of the vorticity in the oncoming flow. It should be noted that only the component of the vorticity which is perpendicular to the flow enters into the relation, and that the magnitude of the effect is proportional to the vane circulation.

Equation [7] was derived under the assumption that there were no entropy gradients in the fluid. In reference (8) it is shown that Equation [7] should contain an additional term if there are (axisymmetric) entropy variations in the flow approaching the blade row. It is shown in reference (11) that this new term is of a smaller order than the others for subsonic boundary-layer secondary flow.

Application at the Hub and Casing. When V_s is zero, the two terms in Equation [7] are equal and opposite. This happens if we consider the sections where the vanes intersect the hub and casing surfaces. Here V_s is forced to be zero since no flow can cross these boundaries. From Equation [7] we get

$$\frac{d\Gamma_{VA}}{dn_1} = -W_1 \zeta_{\perp} \frac{\Gamma_{VA}}{W_{\infty}^2} \dots \dots \dots [8]$$

at the hub and tip sections. Thus, with the flow conditions given upstream, and with the vane circulation specified, Equation [8] gives the gradient of the vane circulation at the hub and tip. For the design problem this is equivalent to the specification of the slope of the tangential velocity distribution at the hub and tip. The restriction caused by Equation [8] does not appear if only axisymmetric flows are considered since for these cases the secondary velocity, as given by Equation [7], is always zero due to the fact that Γ_V is zero with an infinite number of vanes. If, however, it is desired to have the axisymmetric flow capable of being produced as an average by a finite number of vanes, Equation [8] must be considered. The neglect of this consideration becomes more serious if a small number of vanes is used in a blade

row since, for a given amount of turning, the secondary velocity is inversely proportional to the number of vanes. This is immediately evident if Γ_V is expressed in terms of the tangential velocity distributions as is done in the Nomenclature, and then substituted into Equation [7].

It should be noted that when the flow upstream of the blade row is irrotational, the gradient in vane circulation at the hub and tip is zero. This fact was pointed out previously by Siesstrunk and Fabri (12) and Lieblein and Ackley (13).

Some evidence as to the distance from the hub and casing within which the restriction placed by Equation [8] has an effect will be presented later.

THE CHANNEL THEORY

We can think of the passage bounded by two adjacent blades and the hub and casing surfaces as being a channel through which the flow must pass. We consider the channel to be continued upstream and downstream of the blades, with the surfaces composed of stagnation streamlines as its walls. If we were to solve the complete problem, being given only the flow at upstream infinity and the geometry of the blades, the shape of these sheets would be derived. We will assume in this section that the true stagnation streamlines correspond to those obtained from the axisymmetric solution, and on this basis we will attempt to trace the path of a vortex line as it moves through the channel.

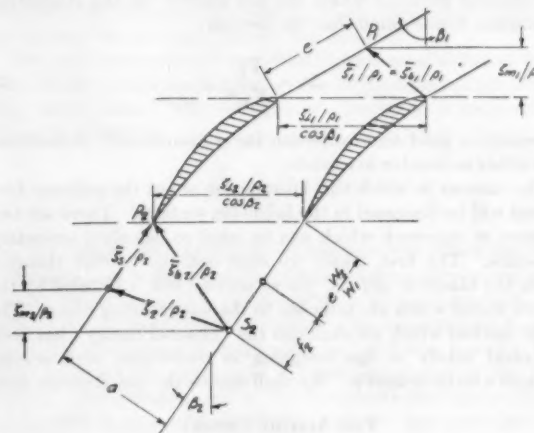


FIG. 3 DETERMINATION OF SECONDARY VORTICITY BY CONSIDERATION OF THE DISPOSITION OF A VORTEX LINE AS IT MOVES THROUGH A BLADE PASSAGE WITH FLUID PARTICLES

Determination of Secondary Vorticity. In Fig. 3 we see a cross section of the channel cut by an axisymmetric stream surface. The subscript b on some of the vorticity vectors indicates that these vectors would be obtained by using the blade-to-blade solution to transport the (absolute) vorticity, this being our best approximation of the actual flow. Vectors without subscripts represent the (absolute) vorticity of the axisymmetric flow.

The scale of the vorticity is chosen such that ζ_1/ρ_1 just fits in between the dividing streamlines, as shown in the figure. Now the Helmholtz laws tell us that as the line of particles tagged by this vector moves through the channel, the vorticity (divided by the density) of the fluid always has the direction of the line of particles and a magnitude proportional to the length of the line. We assume that the line remains straight² so that the vorticity-vector

² In assuming the vortex line straight, we use the average value. It is felt that this is sufficient, in view of some of the other assumptions, and it is certainly desirable, from the point of view of simplification, to have the vorticity constant from blade to blade.

diagrams shown in the figure apply to any point in between the dividing streamlines in front of and behind the system. Since we found previously how much faster a particle travels around one side of the vane than the other, we are able to deduce easily the point S_2 , as shown in Fig. 3. The point which is a distance $\epsilon W_2/W_1$ away from the trailing edge of the vane on the right would mark the tail of the vorticity vectors if particles traveled around the suction and pressure sides of the vanes in the same amount of time. The distance f_2 which is given by

$$f_2 = W_2 \Delta t = W_2 \frac{\Gamma_{VA}}{W_{\infty}^2}$$

accounts for the time difference.

It is not immediately obvious that ξ_2/ρ_2 , which represents the axisymmetric vorticity at station 2, should have the same component perpendicular to the relative flow direction as ξ_1/ρ_1 , as has been indicated in Fig. 3. This can be seen to be true, however, if we note that the circumferential blade pitch represents, to the scale which is being employed here, the quantity

$$\frac{\xi_1/\rho}{\cos \beta}$$

When we consider that the equation governing the axisymmetric flow can be written*

$$\frac{\xi_1/\rho_1}{\cos \beta_1} = \frac{r_2}{r_1} \frac{\xi_2/\rho_2}{\cos \beta_2}$$

we see that this equation is satisfied if ξ_2/ρ_2 fits in between the downstream dividing streamlines, because the ratio of the downstream to upstream blade pitch is just equal to r_2/r_1 . Therefore the three-dimensional and the axisymmetric solutions both have the same perpendicular vorticity component.

The secondary vorticity, which is the difference between the actual vorticity and the axisymmetric vorticity as expressed by the following equation

$$\xi_2/\rho_2 = \xi_{2A}/\rho_2 - \xi_1/\rho_2 \dots \dots \dots [9]$$

is therefore parallel to the relative velocity.

From the geometry of Fig. 3 we see that

$$\xi_2/\rho_2 = f_2 \frac{\xi_1/\rho_1}{a} + \frac{\xi_{m1}/\rho_1}{\cos \beta_1} \frac{W_2}{W_1} - \frac{\xi_{m2}/\rho_2}{\cos \beta_2} \dots \dots [10]$$

This equation is written in units of vorticity over density rather than in units of length. Thus in the first term on the right in Equation [10] we have multiplied the distance f_2 by a scale factor which is seen to change length to vorticity over density. In the second term the length a has already been converted, and so also with the last term. Multiplying by $\rho_2 a$ we find

$$\xi_2 a = f_2 \xi_1 + \xi_{m1} \frac{a}{\cos \beta_1} \frac{\rho_2}{\rho_1} \frac{W_2}{W_1} - \xi_{m2} \frac{a}{\cos \beta_2}$$

Using the facts that

$$\frac{dn_1}{dn_2} = \frac{\rho_2 r_2 V_{m2}}{\rho_1 r_1 V_{m1}}$$

$$\frac{\rho_2}{\rho_1} \frac{W_2}{W_1} = \frac{\rho_2}{\rho_1} \frac{V_{m2} \cos \beta_2}{V_{m1} \cos \beta_1} \frac{r_2}{r_1} = \frac{\cos \beta_1}{\cos \beta_2} \frac{r_1}{r_2} \frac{dn_1}{dn_2}$$

and

$$\frac{a}{\cos \beta_2} = \frac{2\pi r_2}{N}$$

* See, for instance, reference (8).

this becomes

$$\xi_2 a = f_2 \xi_1 + \frac{2\pi}{N} r_1 \xi_{m1} \frac{dn_1}{dn_2} - \frac{2\pi}{N} r_2 \xi_{m2} \dots \dots [11]$$

In order to evaluate this further, we recall that the vane circulation is given by

$$\Gamma_V = \frac{2\pi}{N} (r_1 V_{\theta 1} - r_2 V_{\theta 2})$$

so that the gradient in vane circulation downstream of the blade row is

$$\frac{d\Gamma_V}{dn_2} = \frac{2\pi}{N} \left[\frac{d(r_1 V_{\theta 1})}{dn_1} \frac{dn_1}{dn_2} - \frac{d(r_2 V_{\theta 2})}{dn_2} \right] \dots \dots [12]$$

Since in the axisymmetric flow the meridional vorticity is equal to the normal gradient in angular momentum divided by the radius, Equation [12] can be written

$$\frac{d\Gamma_V}{dn_2} = \frac{2\pi}{N} \left(r_1 \xi_{m1} \frac{dn_1}{dn_2} - r_2 \xi_{m2} \right)$$

When this is substituted into Equation [11] we obtain the result

$$\xi_2 a = f_2 \xi_1 + \frac{d\Gamma_V}{dn_2}$$

Using

$$f_2 = W_2 \frac{\Gamma_{VA}}{W_{\infty}^2}$$

together with the axisymmetric formula

$$W_2 \xi_1 = W_1 \xi_1 \frac{dn_1}{dn_2}$$

we can write

$$\xi_2 a = \left(W_1 \xi_1 \frac{\Gamma_{VA}}{W_{\infty}^2} + \frac{d\Gamma_V}{dn_2} \right) \frac{dn_1}{dn_2} \dots \dots [13]$$

This equation gives us the vorticity which contributes to the secondary flow in terms of quantities which are all known from the primary flow, except Γ_{VA} , which can be replaced by Γ_V for small secondary motions. As with the airfoil theory, it will be found most convenient to regard the vane circulation as always positive and the secondary vorticity positive when it induces secondary velocities on the suction side of the vane which are in the n -direction. ξ_1 is positive when the vector product $\vec{W} \times \xi_1$ points in the n -direction.

Comparison of Airfoil and Channel Theories. The striking similarity between Equation [13] and Equation [7] of the airfoil theory should be noted. The right-hand members of both these equations are identical except for the fact that the actual (primary plus induced) vane circulation is used in Equation [7] whereas the vane circulation used in the second term of Equation [13] is clearly that of the primary flow only.

In order to compare the relations further, we study a strip in the flow between two vortex sheets downstream of the blade row, Fig. 4. We view the strip by looking parallel to the flow in the upstream direction. By Stokes' law, the secondary circulation obtained by integrating the secondary velocity

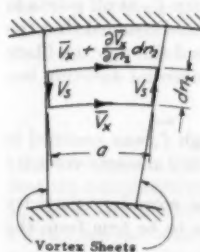


FIG. 4 COMPARISON OF CHANNEL AND AIRFOIL THEORIES BY CONSIDERATION OF VORTICITY ENCOMPASSED BY CIRCUIT IN FLOW AND CIRCULATION AROUND THAT CIRCUIT

around the border of the strip must equal the integral of the secondary vorticity over the area of the strip.

Hence

$$2V_s dn_2 + \bar{V}_s a - \left(\bar{V}_s + \frac{\partial \bar{V}_s}{\partial n_2} dn_2 \right) \left(a + \frac{\partial a}{\partial n_2} dn_2 \right) = \zeta_s a dn_2$$

in which \bar{V}_s indicates the average value of the cross-flow velocity between two vortex sheets. Neglecting higher-order differentials, this becomes, after rearrangement

$$\begin{aligned} 2V_s - \zeta_s a &= \bar{V}_s \frac{\partial a}{\partial n_2} + a \frac{\partial \bar{V}_s}{\partial n_2} \\ &= \frac{\partial}{\partial n_2} (a \bar{V}_s) \dots \dots \dots [14] \end{aligned}$$

With

$$a = \frac{2\pi r_2}{N} \cos \beta_2$$

Equation [14] becomes

$$2V_s - \zeta_s a = \frac{2\pi}{N} \frac{\partial}{\partial n_2} (r_2 \bar{V}_s \cos \beta_2)$$

Now $\bar{V}_s \cos \beta_2$ represents the component of the secondary velocity in the tangential direction, so we may write

$$\begin{aligned} 2V_s - \zeta_s a &= \frac{2\pi}{N} \frac{\partial}{\partial n_2} (r_2 \bar{V}_{\theta s}) \\ &= \frac{d\Gamma_{V_s}}{dn_2} \dots \dots \dots [15] \end{aligned}$$

in which Γ_{V_s} is the increase in vane circulation due to the secondary flow. This equation, which must always hold since it is only a restatement of Stokes' law, relates the secondary velocity in the vortex sheets with the secondary vorticity. We have separately derived relations for both of these quantities, Equation [7] and Equation [13], and these will now be substituted into Equation [15]. Hence we find

$$\frac{d\Gamma_{VA}}{dn_1} \frac{dn_1}{dn_2} - \frac{d\Gamma_V}{dn_1} \frac{dn_1}{dn_2} = \frac{d\Gamma_{V_s}}{dn_2}$$

which is seen to be an identity, this proving that the airfoil theory and the channel theory, as presented in this paper, are compatible.

Determination of Secondary Velocities. By means of Equation [13] we can determine the secondary vorticity ζ_s at all points in the flow downstream of the cascade. We have assumed that the vortex lines remained straight as they passed through the blade channel so that ζ_s does not vary in the tangential direction but only in the normal (spanwise) direction.

It should be pointed out that even though ζ_s was specified in Equation [9] to be the difference between two absolute vorticity vectors, for a rotating system ζ_s is also the secondary vorticity relative to the moving vanes. This is seen to be true from the following relation

$$\zeta_s = \zeta_{s2} - 2\omega = (\zeta_2 - 2\omega)$$

which was obtained from Equation [9] by adding and subtracting 2ω . Thus ζ_s as obtained from Equation [13] can be used in the rotating co-ordinate system.

We establish a surface in the region downstream of the blades by drawing the line $A-A$, Fig. 5, perpendicular to the relative streamlines and midway between the vortex sheets. The lines B_1-C_1 , B_2-C_2 , etc., lie in the axisymmetric stream surfaces and are drawn perpendicular to the streamlines to determine the integration surface.

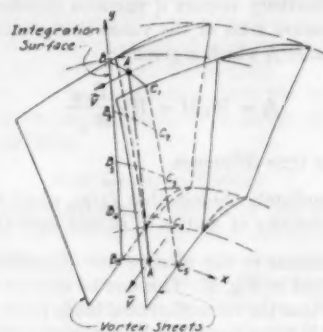


FIG. 5 SPECIFICATION OF SURFACE ON WHICH SECONDARY VELOCITIES ARE TO BE FOUND

As the flow moves downstream the shape of the surface changes because the vortex sheets twist. Since the sheets themselves follow the primary flow, their motion was subtracted from the total motion in arriving at the secondary vorticity. Therefore the secondary motion is of a cellular type in between the sheets. We will only examine the flow at one integration surface downstream of each blade row, and we will assume that the lines $B-C$ intersect the line $A-A$ at right angles for this surface. A more realistic surface could be taken, but the added complication would not be justified in view of the other assumptions that will be made.

In order to make the problem two-dimensional, we assume that we can neglect the fact that the surface is warped. We draw on a plane surface the line $A-A$ with its developed length and then draw the lines B_1-C_1 , B_2-C_2 , etc., at right angles to $A-A$. We retain the component of the curvature of the $B-C$ lines which lies in the integration surface, but we reduce to zero the curvature component normal to the surface, this being a necessary requirement for a plane figure. We use the developed length of these lines.

Let x and y , Fig. 5, together with σ , which points in the direction of flow, be the space variables. The x -direction is always chosen to correspond to the direction of main-stream turning (running from the pressure side of a vane to the suction side of the adjacent vane) and the y -direction corresponds to the n -direction. For the purposes of this paragraph x and y should be considered as being Cartesian co-ordinates so that the boundaries of the surface are not necessarily lines $x = \text{const}$ or $y = \text{const}$. Assuming that the variation in the density of the fluid can be neglected, we can write the continuity equation as follows

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_\sigma}{\partial \sigma} = 0 \dots \dots \dots [16]$$

Here V_x , V_y , and V_σ represent the velocities of the secondary flow. Since we found that the secondary vorticity vector is parallel to the flow direction, the secondary velocities will be induced principally in the x, y -plane. Neglecting any velocities induced in the σ -direction⁷ the foregoing equation becomes

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} = 0 \dots \dots \dots [17]$$

⁷ Further discussion of this point will be given in the Appendix.

The vorticity ζ_s can be expressed as velocity gradients in these co-ordinates

$$\zeta_s = \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \quad [18]$$

Equation [17] allows us to define a stream function ψ , such that

$$V_x = -\frac{\partial \psi}{\partial y} \text{ and } V_y = \frac{\partial \psi}{\partial x} \quad [19]$$

Using these with Equation [18] we obtain Poisson's equation

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = \zeta_s \quad [20]$$

which must be solved for ψ . The boundary conditions are given by the fact that ψ must be constant around the edges of the integration surface in Fig. 5. The velocities deduced from the solution of Equation [20] should be added to the primary flow to obtain the actual flow at the outlet of the blade row.

In order to solve the boundary-value problem at which we have arrived, we expand ψ in a double series of functions each of which satisfies the boundary conditions separately. The nature of these characteristic functions depends on the shape of the boundaries.

One of the principal reasons we are investigating secondary flow is that we wish to establish the amount of induced fluid turning that this flow causes so that we can make corrections for this in the blade-to-blade solution. It is seen that this induced turning is described in terms of V_x . Because this quantity varies across the width of the channel, however, we will want to find its average value. This is desired since the axisymmetric flow, which is the basis of the method outlined in this paper, was shown by Ruden (14) to represent the circumferential average of the flow with a finite number of blades.

In the following paragraphs some results found for the flows in a rectangular surface and a semi-infinite strip will be presented.

Solution With Rectangular Boundaries.

In Fig. 6 we see a rectangular strip of width a and height b with the origin of the co-ordinates at one corner of the strip. Consider the flows induced in such a surface by the family of secondary vorticity distributions represented by

$$\zeta_{si} = k_i \eta^i \quad [21]$$

By letting i assume the values 0, 1, 2, . . . , Equation [21] is seen to represent uni-

form, linear, quadratic, etc., distributions of vorticity. The scale factor k represents the magnitude of the vorticity at the top of the strip ($\eta = 1$).

In order to work with dimensionless quantities, we define

$${}_i V_\xi = \frac{\pi^2}{4} \frac{{}_i V_x}{ak_i} \quad [22]$$

We use the i as a prescript when it would otherwise interfere with the subscripts.

The values for ${}_i V_\xi$ found in reference (8) for several cases are shown in Figs. 7 and 8.

Since Equation [20] is linear in ψ , and since the velocities are obtained from ψ by a linear operation, we can add vorticity distributions in an attempt to approximate the desired distribution ζ_s , and we also can add the velocities deduced for each of the vorticity distributions. Thus if we have

$$\zeta_s(\eta) = \sum_{i=0}^j \zeta_{si}(\eta) \quad [23]$$

it follows that we also have

$$V_s(\eta) = \sum_{i=0}^j {}_i V_x(\eta) \quad [24]$$

In order to use the results plotted in Figs. 7 and 8, we find the constants k_i such that $\zeta_s(\eta)$ is represented by

$$\zeta_s(\eta) = \sum_{i=0}^j k_i \eta^i \quad [25]$$

The total induced velocity would then be, using Equation [22]

$$V_s(\eta) = \frac{4a}{\pi^2} \sum_{i=0}^j k_i {}_i V_\xi(\eta) \quad [26]$$

It is, of course, necessary, when adding the distributions, to be sure that all of them are for the same boundary shape. The upper limit of summation, j , indicates the number of terms taken in the approximation.

In applying the results of this section to practical cases in turbomachinery we must make some allowance for the fact that the width of the strip a is not constant from hub to tip. For lack of a better method, it is suggested that the rectangular strip whose value of a is nearest to that of the actual strip be used, but that the correct values of a be used in Equation [26] for each value of η . If the shape of the strip is far from rectangular, and does not resemble any other geometric shape for which characteristic functions are known, the problem is best approached through the solution of Equation [20] by relaxation methods.

Solution for a Semi-Infinite Strip. If there is a large concentration of vorticity near one end of the integration surface of Fig. 5 such as that which would be caused by a boundary layer on the hub or casing in the upstream flow, we can replace the surface by a semi-infinite strip. In Fig. 9 we see such a strip of width a . Consider the flows induced in such a strip by the family of secondary vorticity distributions represented by

$$\zeta_{sm}(\eta) = \zeta_0 e^{-m\eta\eta_1} \quad [27]$$

These represent distributions of vorticity which have the values ζ_0 at the bottom of the strip and which drop off to zero exponentially with η . By adjusting the constant m the rate at which the function diminishes can be controlled. Results for three values of m have been plotted in Fig. 10. As before, we write the velocities in a nondimensional form such that the actual dimensional values are obtained by multiplying the nondimensional values (indicated by Greek subscripts) by

$$\frac{4}{\pi^2} a \zeta_0$$

It is interesting to note for the case of uniform vorticity ($m = 0$) that the x -averaged secondary velocity is largest at the bottom of the strip and diminishes to a negligible value within one strip width from the bottom. Thus we would expect that the restriction discussed earlier in connection with Equation [8] is limited in effect to a region extending no further than one blade spacing from the hub or casing.

The vorticity distribution corresponding to $m = 4$ is seen to resemble that which might be caused by a boundary layer upstream of the cascade. The induced velocity ${}_4 V_\xi$ is such that it would cause overturning near the wall but underturning a short distance away from the wall and up into the main stream. This phe-

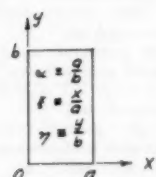


FIG. 6 RECTANGULAR STRIP

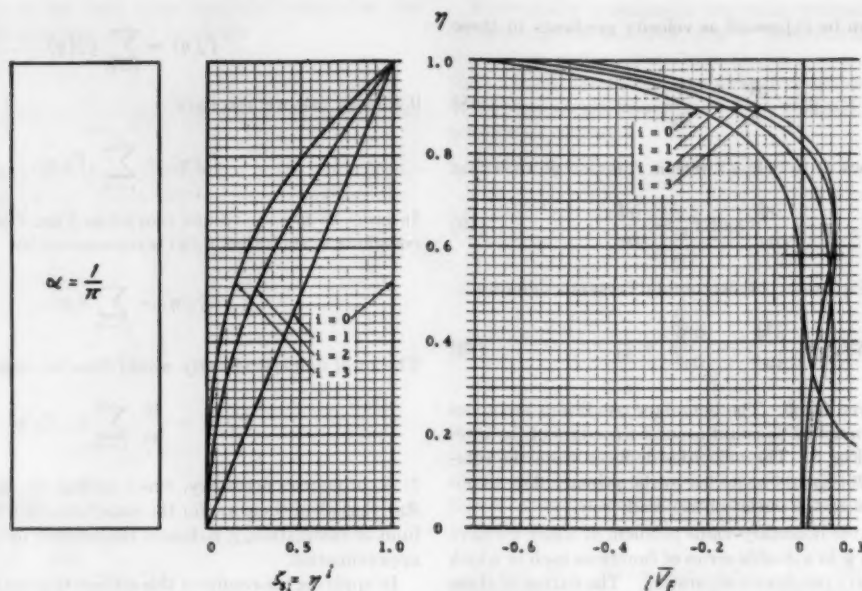


FIG. 7 BLADE-TO-BLADE AVERAGED CROSSFLOW VELOCITIES FOR $\xi_i = \eta^i$, ($i = 0, 1, 2, 3$), $\alpha = 1/\pi$

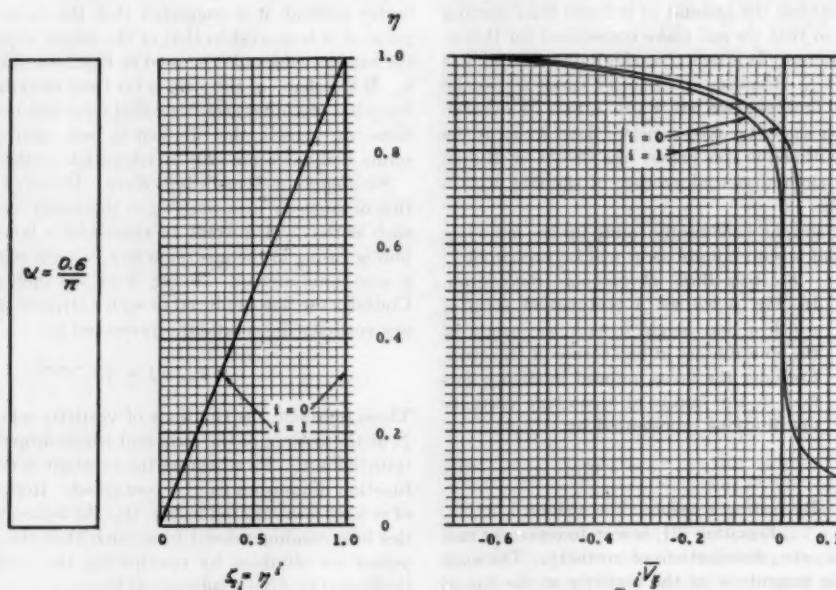


FIG. 8 BLADE-TO-BLADE AVERAGED CROSSFLOW VELOCITIES FOR $\xi_i = \eta^i$, ($i = 0, 1$), $\alpha = 0.6/\pi$

nomenon has been observed experimentally by many investigators. Because of the large stream-surface distortions which occur with boundary layers, the method of this paper will, of course, give only qualitative information about such flows. Some work on the problem of large distortions has been done by Loos (15).

CONCLUSIONS

1 When the flow approaching a vane or wing is rotational and steady, the spanwise velocities which make up the trailing-vortex sheet shed from the vane are caused by two effects, one being the well-known gradient in vane-circulation effect, which is present even with irrotational approach flow, and the other being caused by the distributed vorticity in the flow.

2 Because the spanwise velocities which make up the trailing-vortex sheet must be zero at the hub and casing surfaces, a relation is found which gives the gradient in vane circulation in terms of the upstream flow and the vane circulation itself at these locations. For the special case of irrotational upstream flow, this relation gives the result previously reported by Siestrunk and Fabri (12) and by Lieblein and Ackley (13) that the vane-circulation gradient must be zero at the hub and casing. The effects of this necessary condition at the walls extend into the stream a distance roughly equal to the distance a between stagnation streamlines.

3 The secondary flow was found to be proportional to the vane circulation (for given velocity triangles) and hence inversely proportional to the number of vanes in a row. For numbers of blades

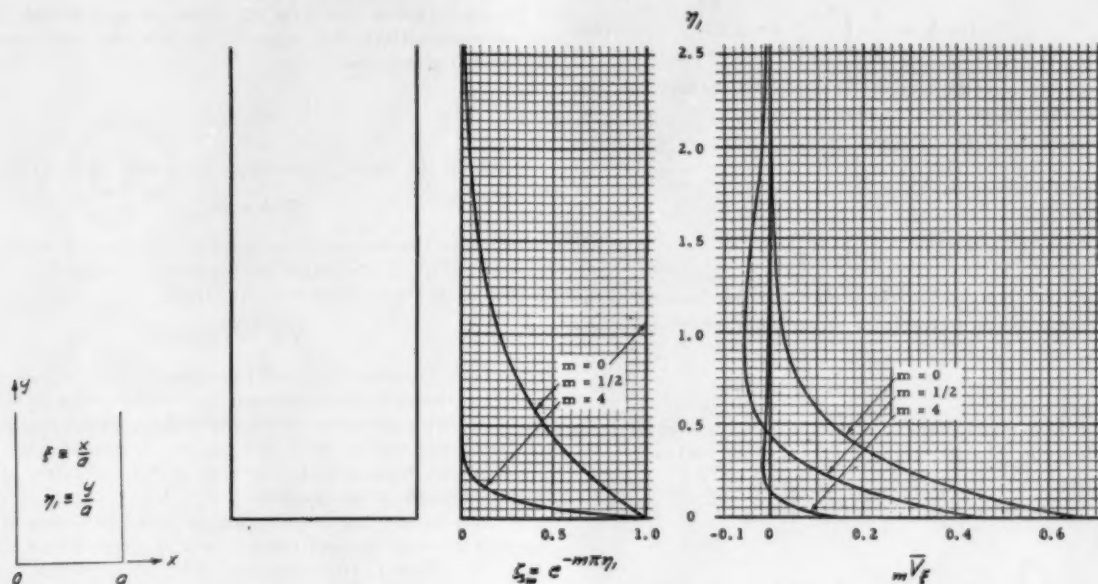


FIG. 9 SEMI-INFINITE STRIP

FIG. 10 BLADE-TO-BLADE AVERAGED CROSSFLOW VELOCITIES FOR $\xi = e^{-m\pi y/a}$, ($m = 0, 1/2, 4$)

which are commonly used in practice in the jet-engine field, the secondary flow is usually small enough to be neglected.

4 A good estimate of the magnitude of secondary flow to be expected can be easily obtained from the axisymmetric solution and an assumed number of vanes. To do this, it is only necessary to calculate Γ_f from the formula given in Nomenclature, and then to find V_s from Equation [7]. The number of vanes can then be adjusted to give acceptably small secondary flow.

BIBLIOGRAPHY

- 1 "Two-Dimensional Flow on General Surfaces of Revolution in Turbomachines," by J. D. Stanitz and G. O. Ellis, NACA TN 2654, March, 1952.
- 2 "An Approximate Method of Determining the Subsonic Flow in an Arbitrary Stream Filament of Revolution Cut by Arbitrary Turbomachine Blades," by C-H. Wu, C. A. Brown, and V. D. Prian, NACA TN 2702, June, 1952.
- 3 "The Secondary Flow in a Cascade of Airfoils in a Non-Uniform Stream," by H. B. Squire and K. G. Winter, *Journal of the Aeronautical Sciences*, vol. 18, 1951, pp. 271-277.
- 4 "Secondary Circulation in Fluid Flow," by W. R. Hawthorne, Gas Turbine Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., May, 1950.
- 5 "Secondary Flows in Fluid Dynamics," by R. E. Kronauer, Pratt and Whitney Research Report No. 132, Gordon McKay Laboratory, Department of Engineering Sciences and Applied Physics, Harvard University, Cambridge, Mass., April, 1951.
- 6 "Shear Flow in Bends," by H. P. Eichenberger, Gas Turbine Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., April, 1952.
- 7 "Secondary Flow in Cascades of Twisted Blades," by F. F. Ehrich, *Journal of the Aeronautical Sciences*, vol. 22, 1955, pp. 51-60.
- 8 "Three-Dimensional Flow in Axial-Flow Turbomachinery, Part I, Theoretical Determination of Secondary Flow," by L. H. Smith, Jr., Report I-14, Inst. for Coop. Res., Mechanical Engineering Department, The Johns Hopkins University, Baltimore, Md., November, 1953.
- 9 "Three-Dimensional Flow in Axial-Flow Turbomachinery, Part II, Experimental Investigations," by L. H. Smith, Jr., Report I-16, Inst. for Coop. Res., Mechanical Engineering Department, The Johns Hopkins University, Baltimore, Md., November, 1953.
- 10 "On Rotational Gas Flows," by A. Vazsonyi, *Quarterly of Applied Mathematics*, vol. 3, 1945, pp. 29-37.
- 11 "A Note on the Effect of Entropy Gradients in the Analysis of the Flow Through Turbomachines," by L. H. Smith, Jr., May, 1954, unpublished.
- 12 "Ecoulements tourbillonnaires dans les machines axiales,"

translation—"Vortex Flow in Axial Machines," by R. Siestrunk and J. Fabri, ONERA Pub. No. 45, 1950, NACA Translation No. 29598.

13 "Secondary Flows in Annular Cascades and Effects on Flow in Inlet Guide Vanes," by S. Lieblein and R. H. Ackley, NACA RM E51G27, August, 1951.

14 "Investigation of Single Stage Axial Fans," by P. Ruden, NACA TM 1062, April, 1944.

15 "Analysis of Secondary Flow in the Stator of an Axial Turbomachine," by H. G. Loos, Daniel and Florence Guggenheim Jet Propulsion Center, California Institute of Technology, Pasadena, Calif., Technical Report No. 3, Contract AF-18(600)-178, September, 1953.

Appendix

The primary flow which we chose to transport the vorticity has no velocity component normal to the axisymmetric stream surfaces, so that the vortex lines remain in these surfaces as they move through the blade rows. Actually, because of the secondary motion, the flow surfaces which were axisymmetric are twisted through an angle in the x, y -plane, and the secondary vorticity acquires a component normal to the axisymmetric stream surfaces which we shall call ξ_s . Because of ξ_s there is flow induced in the σ -direction (the direction of the primary flow).

Let the angle in the x, y -plane through which a stream surface twists because of secondary flow be λ . It is assumed that the secondary flow is linear from blade to blade so that λ is constant in the x -direction. Let $V_{s\sigma}$ be the secondary velocity in the normal direction of a particle which travels near the suction wall of the channel, that is, near the blade suction surface and, consequently, near the vortex sheet downstream of the blade. Then

$$\tan \lambda = \frac{1}{a/2} \int_0^{\tau} V_{s\sigma} dt$$

in which τ is the time taken for a particle to travel from the inlet of the blade row to the point under consideration. We can change this from a time integral to a distance integral with the substitution

$$d\sigma = W dt$$

where σ is the distance along a streamline measured from the inlet of the blade row and W is the primary velocity. Thus

$$\tan \lambda = \frac{2}{a} \int_0^a \frac{V_{ys}}{W} d\sigma \dots \dots \dots [28]$$

Because the absolute vorticity is carried by the flow particles in the surface which twists, we find

$$\zeta_n = \zeta_{\perp} \tan \lambda$$

and, from Equation [28]

$$\zeta_n = \zeta_{\perp} \frac{2}{a} \int_0^a \frac{V_{ys}}{W} d\sigma \dots \dots \dots [29]$$

If we assume that ζ_n does not induce any flow normal to the vortex sheets, which is reasonable for our physical problem, then

$$V_{\sigma} = \left(\frac{a}{2} - x \right) \zeta_n$$

in which V_{σ} represents the secondary velocity induced by ζ_n in the σ -direction. Joining this with Equation [29] we have

$$V_{\sigma} = \left(\frac{a}{2} - x \right) \zeta_{\perp} \frac{2}{a} \int_0^a \frac{V_{ys}}{W} d\sigma$$

and

$$\frac{\partial V_{\sigma}}{\partial \sigma} = \left(\frac{a}{2} - x \right) \zeta_{\perp} \frac{2}{a} \frac{V_{ys}}{W}$$

We have already assumed that V_y varies linearly from blade to blade, that is

$$V_y = - \left(\frac{a}{2} - x \right) \frac{2}{a} V_{ys}$$

so that we may write

$$\frac{\partial V_{\sigma}}{\partial \sigma} = - \frac{V_y}{W} \zeta_{\perp}$$

Kronauer (5) arrived at this result by considering that the primary velocity of a particle remains constant as it moves in the y -direction.

The complete continuity equation can now be seen to be

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} = \frac{V_y}{W} \zeta_{\perp} \dots \dots \dots [30]$$

which should be used instead of Equation [17]. Let us try to assess the error which arises when Equation [17] is used. Equation [30] describes the flow in a two-dimensional region with distributed sources of strength $(V_y/W)\zeta_{\perp}$ per unit area. We can think of the complete secondary flow as being composed of two parts superimposed, one part (which we already have solved) having distributed vorticity but no sources, and the other part being irrotational but with distributed sources. Now a difficult point arises here in that the source strength is not known until both parts have been solved, because the source strength is a function of the complete V_y . However, if the contribution to V_y from the source flow solution is small compared to that from the vorticity flow distribution, it is permissible to neglect the former. If this is not the case, we must iterate using the sum of the two V_y 's to find the new source distribution, or we must abandon the idea of solving two relatively simple flows separately and attack the problem as a whole.⁸

⁸ Kronauer (5) has done this and arrived at a second-order linear differential equation with V_y as the dependent variable. Judging from a numerical example which he gives, some tedium is involved in satisfying the boundary conditions by trial-and-error numerical solution, but in some cases this may still be easier than iterating as suggested above.

In order to get an idea of the source flow strength so that it can be compared with the flow caused by the vorticity, we can define a potential ϕ such that

$$\vec{V}' = -\nabla \phi$$

where \vec{V}' is the velocity induced by the source field. Then

$$\nabla^2 \phi = S \dots \dots \dots [31]$$

in which S is the strength of the distributed sources. For our case $S = (V_y/W)\zeta_{\perp}$. The part of the complete secondary flow which is induced by the vorticity is represented by

$$\nabla^2 \psi = \zeta_{\perp} \dots \dots \dots [20]$$

Since both Equation [31] and Equation [20] are of the same form, and since velocities are found from both ϕ and ψ by taking first derivatives, and since, furthermore, the boundary conditions on both equations are such that no flow is induced when the equations are homogeneous, the ratio S/ζ_{\perp} is indicative of the relative strength of the two flows.

We thereby arrive at the following procedure for testing to find the significance of the source flow. First, ζ_{\perp} is found from Equation [13]. Then V_s (the maximum of V_y at any radius) is estimated from Equation [7]. A maximum value of S is then calculated from $S = (V_s/W)\zeta_{\perp}$ and the ratio S/ζ_{\perp} is examined. If S/ζ_{\perp} is less than 0.2, we can neglect the source solution. If this ratio has a value higher than 0.2, we should include the source effect or use a different method.

It has been found in reference (9) that, for secondary flow other than that caused by boundary layers, the source effect usually can be neglected.

Discussion

H. G. Loos.⁹ In this work the author has succeeded in making an important contribution toward our understanding of rotational flows past airfoils. In particular, it is fortunate he found that the strength of the trailing-vortex sheet behind an airfoil in nonuniform flow is not equal to the spanwise derivative of the circulation, but that there is an additional contribution due to the vorticity in the incoming flow. One might gain more physical insight in this matter by investigating the transport of the vortex lines with the flow past an airfoil element in a cylindrical surface, which is coaxial with the blade row. Because of the different velocities at suction and pressure surface, the "spacing" of the vortex lines must be larger at the suction surface than at the pressure surface, so that there is a difference in the vorticity flux through these surfaces. This amount has to be shed in the trailing-vortex sheet and represents the additional contribution due to the vorticity in the incoming flow.

In the Appendix the influence of the twist of the surfaces of constant total pressure is calculated. This linearized analysis is only valid for small secondary vorticity so that in Equation [30] V_x and V_y are of order ζ_{\perp} , and the right-hand side is of order $(\zeta_{\perp})^2$. Hence the twisting gives only a second-order contribution to the secondary velocity. It does not seem justified that this second-order effect is accounted for while another effect of the same order (induced transportation and deformation of the secondary vorticity) has been neglected.

Equation [30] also holds for a flow with a large nonuniformity, which is slightly turned or disturbed.¹⁰ In this case Equation [30]

⁹ California Institute of Technology, Pasadena, Calif.

¹⁰ "Lifting Line Theory for a Wing in Non-Uniform Flow," by Th. von Karman and H. S. Tsien, *Quarterly of Applied Mathematics*, vol. 3, 1945, p. 1.

has significance because, although V_x/W and V_y/W are small, ζ_\perp is not small so that the right-hand side is now small of first order.

G. L. MELLOR, JR.,¹¹ AND R. C. DEAN, JR.¹² This paper which is a condensation of a larger report (author's reference 8), together with the author's experimental work (author's reference 9), represents a major contribution to the theory and understanding of secondary flow in axial turbomachinery with finite blade spacing.

The statement in the author's conclusion 4, that "the number of vanes can be adjusted to acceptably small secondary flow," since the secondary velocity is inversely proportional to the number of blades, naturally leads us to the question—how much secondary flow is "acceptably small?"—or, in general, what effect does secondary flow have on axial compressors?

This is answered in part by the author when he points out that for a conventional compressor design, secondary flow has small effect on the prediction of air angles through axisymmetric solutions. This fact has been experimentally verified by the author (reference 9), Bowen, Sabersky, and Rannie,¹³ and Beatty, Savage, and Emery.¹⁴

We should now inquire whether or not secondary flow leads to losses. The kinetic energy of the secondary flow itself may be shown to be very small¹⁵ and its eventual dissipation insignificant. However, in a compressor stage which is designed for an entering flow with a radial stagnation-pressure gradient, secondary flow or vortex sheets (they do not exist separately) have a chance to bring about significant losses.

For a basis of discussion, let us consider a "typical" compressor stage where the outlet-velocity distribution is identical or almost identical with the inlet distribution. For this case the stagnation-pressure gradient will not vary from blade row to blade row. Furthermore, the work distribution must be constant with radius to satisfy this condition. If we assume that the entering flow is axisymmetric, which is true on a time average, then properties in the tangential direction will be constant and surfaces of constant stagnation pressure will be surfaces of revolution before the blade row. If the blade design is such that secondary velocities exist, then stagnation pressure-surfaces will warp in passage through the blade row. The outlet flow will exhibit skewed surfaces with the result that a component of the stagnation-pressure gradient will be in the tangential direction. The fluid eventually will mix out to a constant stagnation pressure at any given radius. The mixing process, of course, involves losses, the amount of which is at present open to speculation but it may be appreciable.

It is felt by some investigators that losses due to time fluctuations in an entering relative flow may be significant. The fluctuations are usually considered to be due to preceding blade wakes. However, a "sawtoothed" velocity distribution in the tangential direction will be present in the exit flow just described. This should increase losses due to time fluctuations.

From the definition of Γ_{VA} in the author's nomenclature, it can be seen that $d\Gamma_{VA}/dn = 0$ for the constant work stage under discussion. Also a now common fluid-dynamic relationship is

¹¹ Gas Turbine Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

¹² Gas Turbine Laboratory, Massachusetts Institute of Technology, Cambridge, Mass. Assoc. Mem. ASME.

¹³ "Investigation of Axial Flow Compressors," by J. P. Bowen, R. H. Sabersky, and W. D. Rannie, *Trans. ASME*, vol. 73, 1951, pp. 1-15.

¹⁴ "Experimental Investigation of Flow Through Three Highly Loaded Inlet Guide Vanes Having Different Spanwise Circulation Gradients," by L. A. Beatty, Melvyn Savage, and J. C. Emery, NACA RM L52D25a, July 28, 1952.

¹⁵ "Vorticity and Vortex Sheets in Axial Turbomachinery," by G. L. Mellor, Jr., SM thesis, Massachusetts Institute of Technology, Cambridge, Mass., 1954.

$WS_\perp = \frac{1}{\rho} \frac{dP_0}{dn}$ the stagnation-pressure gradient. If we insert these terms into the author's Equation [7] we see that

$$2V_s = (\text{vortex sheet strength}) \approx \frac{1}{\rho} \frac{dP_0}{dn} \frac{\Gamma_{VA}}{W_\infty^2}$$

where the vortex sheet strength is proportional to the radial stagnation pressure gradient.

Our conclusion is that along with the design of "typical" compressor stages with varying stagnation pressure, there are inherent losses which are chiefly a function of the stagnation-pressure gradient itself. If these losses are significant, then this is the importance of secondary flow.

G. F. WISLICENUS.¹⁶ The writer shares with others at the Mechanical Engineering Dept. of the Johns Hopkins University a certain pride about this paper as a good job done by a member of their team.

The paper is a condensed version of two reports (references 8 and 9 of the paper) on an extensive investigation by the author. Whereas the most essential theoretical developments are presented quite clearly in this paper, there was no room nor time left to include the results of concurrent experimental investigations by which the author was able to confirm, at least in principle, the practical validity of his theoretical findings. The writer feels the author should add a brief account of the comparisons between his theoretical and experimental results, as well as of the practical means by which this comparison was achieved. It also will be of interest that the test stage which the author designed and built for this investigation turned out to be an efficient and promising element of an axial compressor which performed essentially in accordance with the author's design expectations. Finally, the writer believes that the author's simple method of observing the three-dimensional flow conditions after a stationary blade row are well worth mentioning.

AUTHOR'S CLOSURE

The author wishes to thank the discussers for their contributions to the paper.

As Dr. Loos suggests, the formula for the vortex-sheet strength, Equation [7], can be derived by considering the transport of vortex lines past an airfoil. I have done something similar to this in reference (8), but the idea of conservation of vorticity flux into the blade is somewhat different and is a convincing way to look at the concept.

The linearized analysis of the paper is only valid for small secondary vorticity, small secondary velocities, and small angles of stream surface distortion, λ . In short, the secondary flow must represent a small perturbation to the primary flow. It does not follow, however, that the vorticity, ζ_\perp of the primary flow itself has to be small. Thus the case mentioned by Dr. Loos in the last paragraph of his discussion is an admissible one and justifies the inclusion of the Appendix. If the primary flow with large ζ_\perp is more than "slightly" turned or disturbed, the analysis breaks down because the "induced transportation and deformation of the secondary vorticity" mentioned by Dr. Loos becomes significant. Unfortunately, most of the interesting cases of boundary-layer secondary flow fall into this category.

Mr. Mellor and Professor Dean have pointed out one of the consequences of secondary flow in a turbomachine that has a radial stagnation-pressure gradient, i.e., a sawtoothed tangential gradient will be formed. Such a gradient was found in a machine that the author tested; the results are presented in reference

¹⁶ Director of Water Tunnel and Professor, Garfield Thomas Water Tunnel Division, Ordnance Research Laboratory, Pennsylvania State University, University Park, Pa. Mem. ASME.

(9). At a distance of five blade chords downstream of the last blade row in the machine (a stator), the corners of the sawteeth had been rounded off giving a sine-wavelike distribution. Accurate loss measurements were not obtained.

The experiments mentioned by Professor Wislicenus were conducted in a 36-in.-diam low Mach number compressor-test stage consisting of inlet guide vanes, rotor, and stator. The stage was designed to produce highly rotational flows and to have radially nonconstant energy addition. The secondary flows caused by these effects were given primary attention in the experiments; the hub and casing boundary layer secondary flows were examined in less detail.

Since it was impossible to measure the induced turning \bar{V}_x with any degree of accuracy, the vortex sheets that were shed from the blades were examined. The following technique was used: The machine was provided with windows at each station. A flag consisting of a piece of knitting wool about 1 in. long was fastened to a sewing needle with a piece of fine thread so that the needle exerted no moments on the flag. The needle was affixed to a probe stem and thereby the flag was suspended in the flow near the trailing edge of a blade. As the flag was slowly moved in the tangential direction behind the blade, an observer viewed it from the other side of the machine along a line of sight coinciding with the tangential direction through a sighting tube equipped with cross hairs. The quantity that was measured was the pitch angle φ that a meridional streamline makes with the axis of the machine.

The results of one of the traverses behind a stator blade is shown in Fig. 11. A jump in meridional flow angle of about 5 deg was found across the blade wake at that location, and from this the

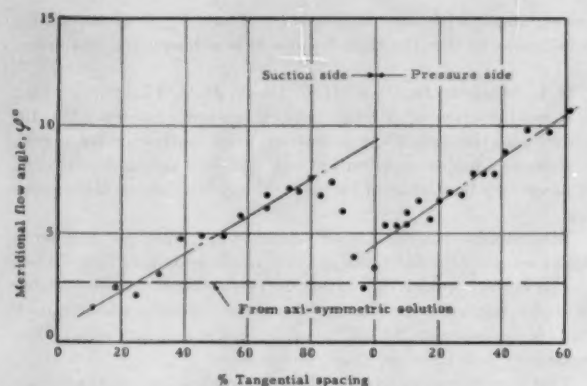


FIG. 11 MEASURED VALUES OF MERIDIONAL FLOW ANGLE BEHIND STATOR BLADES NEAR MID-SPAN RADIUS

vortex-sheet strength was calculated. The continual increase of φ from the left to the right in the figure over and above the periodic sawtooth effect was caused by the fact that the flag could not be viewed in a purely tangential direction because of limited window area. This discrepancy was accounted for in the analysis of the data. The vortex-sheet strengths deduced in this manner checked Equation [7] to within experimental accuracy, but since derivatives of measured quantities had to be taken, a precise comparison could not be made. No secondary-flow measurements were obtained for the rotor so the theory remains unchecked for rotating blade rows.

Development of a Miniature Electrohydraulic Actuator

By S.-Y. LEE¹ AND J. L. SHEARER,² CAMBRIDGE, MASS.

A qualitative discussion is given of the factors establishing the need for a compact electrohydraulic actuator to operate at a relatively low input-power level and capable of providing an output motion proportional to the magnitude of an electric input signal. The basic requirements of such an actuator are discussed, and then the design of each of its components including torque motor, valve, ram, and feedback mechanism, are described in detail. The results of a dynamic analysis of the complete system are employed to show how adequate system damping is provided, and to show how theoretical speed of response is obtained. The results of steady-state and dynamic tests are compared with the results of the analytical work. This actuator has been found to fulfill the needs of a number of instrumentation and control applications, and preliminary tests using compressed air instead of hydraulic oil indicate that pneumatic operation of such a device is also feasible.

NOMENCLATURE

The following nomenclature is used in the paper:

- a = length of valve-plate suspension arms, in. (0.750)
- a_0 }
 a_1 } = coefficients of characteristic equation of actuator
 a_2 }
- A = ram area, sq in. (0.025)
- A_o = orifice area, sq in.
- b = viscous-damping coefficient, lb-sec/in.
- b_1 = height of outer legs of torque-motor field magnet, in. (0.125)
- c = half-height of center leg, in. (0.063)
- C_1 = valve coefficient, in⁴/lb-sec
- C_2 = ram leakage coefficient, in⁴/lb-sec
- C_d = orifice discharge coefficient (0.625)
- d = width of all magnet legs, in. (0.250)
- D = derivative with respect to time, d/dt
- F_1 = force at air gap No. 1, lb
- F_2 = force at air gap No. 2, lb
- F_f = friction force on valve plate, lb
- F_l = external load force, lb
- F_p = pressure force on bottom of valve plate, lb (20)
- F_{rs} = friction force at ram seals, lb
- F_s = feedback-spring force, lb
- F_{sf} = steady-flow force, lb
- F_t = torque-motor force, lb

- F_{us} = unsteady-flow force, lb
- F_{vd} = viscous-damper force, lb
- F_{vs} = valve-spring force, lb
- g = thickness of center-leg air gap of torque motor, in. (0.0015)
- g_1 = thickness of air gap No. 1, in.
- g_2 = thickness of air gap No. 2, in.
- g_0 = thickness of g_1 and g_2 when they are equal, in. (0.0045)
- I_1 = electric current in coil No. 1, amp
- I_2 = electric current in coil No. 2, amp
- k = composite spring stiffness associated with torque motor and suspension valve including stiffness of feedback spring, lb/in. (73.7)
- k_1 = valve coefficient, sq in/sec (168.0)
- k_2 = $C_1 + C_2$, in⁴/lb-sec (4×10^{-4})
- k_3 = fluid-compliance coefficient, in⁴/lb
- k_4 = electromagnetic constant, lb-in²/weber²
- k_5 = electromagnetic constant, webers/amp-turn-in.
- k_a = torque-motor coefficient, lb/amp (11.1)
- k_b = torque-motor coefficient, lb/in. (20.7)
- k_s = stiffness of feedback spring, lb/in. (29.6)
- k_{sf} = steady-flow-force coefficient, lb/in. (80.0)
- k_{rp} = effective stiffness of torque-motor-reed pivot, lb/in. (43.0)
- k_{us} = unsteady-flow-force coefficient, lb-sec/in.
- k_v = stiffness of valve suspension, lb/in. (11.8)
- m = net effective mass of valve plate and moving parts of torque motor, lb-sec²/in. (6×10^{-4})
- m_1 = mass load attached to ram, lb-sec²/in. (1.29×10^{-3})
- n = number of turns per coil (2000)
- P = pressure drop across orifice, psi
- P_m = pressure difference across ram, psi
- P_s = supply pressure, psi (2000)
- Q = flow rate, cu in/sec
- Q_l = leakage flow rate past ram, cu in/sec
- r = radius of torque-motor armature, in. (0.375)
- V = volume of fluid under compression in one end of ram, cu in. (0.0025)
- X = valve position, in.
- Y = ram position, in.
- β = bulk modulus of hydraulic fluid, psi (2.5×10^6)
- Δ = small change
- ρ = mass density of fluid, lb-sec²/in.⁴ (8×10^{-4})
- τ = time constant, sec (0.0002)
- ϕ_1 = magnetic flux in air gap No. 1, webers
- ϕ_2 = magnetic flux in air gap No. 2, webers
- ω_n = natural frequency, rad/sec (3500)
- ω_{ms} = natural frequency of ram and mass load, rad/sec (8870)
- i = subscript denoting initial condition
- ss = subscript denoting steady state

¹ Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology.

² Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology. Mem. ASME.

Contributed by the Instruments and Regulators Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 1, 1954. Paper No. 54-A-196.

INTRODUCTION

In many high-performance control systems for industrial and military applications it is necessary to create a mechanical displacement (e.g., move a shaft) against a variable load such that the displacement is proportional to the magnitude of a low-energy-level electric signal.

Where a high-power-level output motion with fast response to changes of input signal is required, an electromechanical transducer and two or more stages of hydraulic amplification (1)² may be needed. Experience gained in the Dynamic Analysis and Control Laboratory at the Massachusetts Institute of Technology with the development of torque motors (2, 3, 4) indicates that speed of response of a torque motor decreases rapidly when it is designed for increased output-power (stroke-and-force) level. At the same time, torque-motor size grows rapidly with increased output-power level. Such results seem to indicate that electromechanical transducers can be used most effectively at the input with very low electric-power levels followed by a first stage of hydraulic amplification taking place at a relatively low power level. An electrohydraulic actuator composed of a very small torque motor, control valve, ram, and feedback mechanism has been developed which is very compact and which is by itself, without further amplification, directly useful in many instrumentation and control applications.

DESIGN REQUIREMENTS

The actuator discussed in this paper was designed specifically for stroking a 5-hp 4-way valve. The design requirements are as follows:

Stroke.....	± 0.015 in.
Peak output force.....	50 lb
Input signal power level.....	less than 5 watts
Operating pressure.....	2000 psi
Load mass (weight).....	1 lb
Peak output velocity.....	18 ips
Speed of response.....	90° phase lag at 200 cps

A linear relationship between output position and electric-input signal (with no load on output) was required within ± 1 per cent. Hysteresis, the ratio h/Y_{\max} in Fig. 1, was not to exceed 1 per cent during a test made with full-amplitude output motion. Load sensitivity as measured by the magnitude of output motion caused solely by the application of a 35-lb load force was not to exceed 5 per cent of maximum output motion.

Of utmost importance was the need for reliability of operation with emphasis on simplicity of design and ease of application. Operation with hydraulic fluid slightly contaminated by foreign particles was essential, but incorporation of a very small filter in the actuator was permissible.

DESCRIPTION OF MINIATURE ACTUATOR

Fig. 2 is a schematic diagram of the complete actuator system. A small 4-way control valve with a hole-slot-and-plug construction (5) is used to modulate the flow of hydraulic fluid from a constant-pressure source to the chambers of a double-acting ram which delivers the power required to produce output motion. The control valve moves in response to various forces acting on it. These forces consist of the electromagnetic force generated by the flow of electric current (input signal) through the torque-motor coils and the force exerted by the feedback spring, the lower end of which is attached to the ram. The feedback is negative. Initially, the torque-motor current is steady, the valve is centered, and the ram is motionless at a position such that the net force due to deflection of the feedback spring just balances the electromagnetic force generated by the flow of current in the torque-motor coils. A change in torque-motor current may cause leftward valve motion which produces a flow tending to move the ram to the right, thereby increasing the spring deflection and creating a force acting in opposition to that generated by the torque motor; hence, if the system is dynamically stable, the valve eventually will be returned to its center position, and the

ram will assume a new position which is proportional to the magnitude of the torque-motor current. Fig. 3 is a functional block diagram of the system showing the sequence of events just described.

Although the series of events in the preceding description starts with operation of the torque motor and although the valve and torque-motor designs are intricately interrelated, the control valve is in many respects the heart of the system.

A photograph of the complete actuator is shown in Fig. 4.

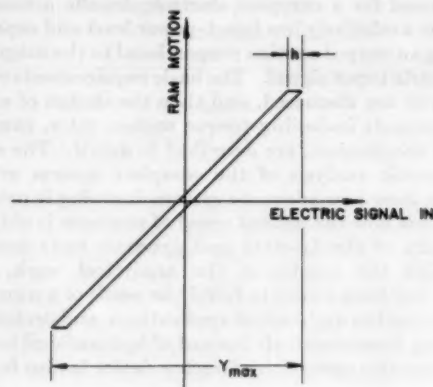


FIG. 1 PLOT OF OUTPUT POSITION VERSUS INPUT SIGNAL AT VERY LOW FREQUENCY

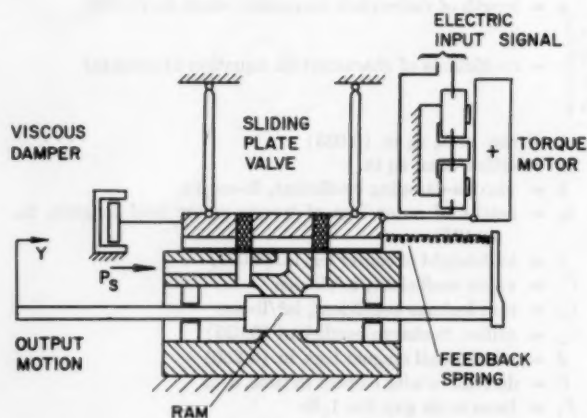


FIG. 2 SCHEMATIC DIAGRAM OF MINIATURE ELECTROHYDRAULIC ACTUATOR

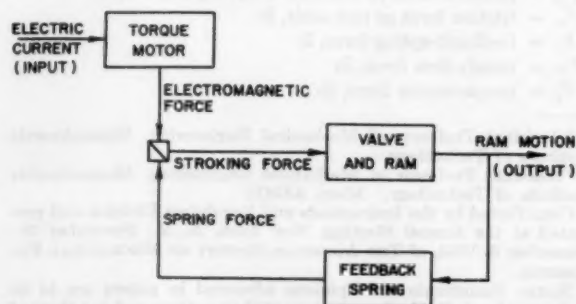


FIG. 3 FUNCTIONAL BLOCK DIAGRAM OF MINIATURE ACTUATOR SYSTEM

² Numbers in parentheses refer to the Bibliography at the end of the paper.

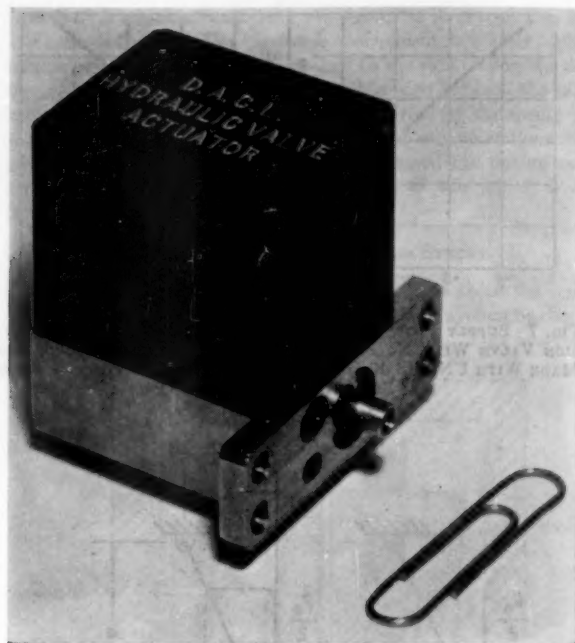


FIG. 4 MINIATURE ACTUATOR COMPARED TO ORDINARY PAPER CLIP

DESIGN CHARACTERISTICS OF THE COMPONENTS

Control Valve. Because of the decision to use a low-power-level torque motor with a small stroke, it is necessary to employ a valve which will provide the maximum required ram velocity with small valve displacement. In order to minimize quiescent leakage and meet the load-sensitivity requirement, a closed-center valve (zero overlap) with small clearance between fixed and moving parts is required. Such factors make it necessary to achieve very accurate relative location of the corners of the metering orifices in the fixed and moving parts of the valve. Friction forces in the valve must be minimized in order to keep torque-motor size down. From an economic point of view, ease of manufacture is an important factor, since cost inevitably must play an important role in the acceptance of the actuator for general use.

Fig. 5 is an isometric drawing of the flexure-suspended valve-plate construction which was developed for the actuator. This unit measures approximately $\frac{1}{2}$ in. \times $\frac{3}{4}$ in. \times $\frac{3}{4}$ in., and it may be machined from a solid piece of steel or fabricated from simple parts. The thin portions of the vertical suspension arms serve as flexure pivots which enable the valve plate to move freely along only its major axis while the upper ends of the arms are held motionless by the rest of the structure which is firmly fastened at the bottom to a block which serves as the fixed part of the valve. A similar type of suspension employing a construction with the suspension arms in tension rather than in compression proved to be more difficult to manufacture and adjust.

A nominal clearance of 0.0002 in. between the valve plate and valve body is attained during final grinding of the bottom surface of the suspension unit. Very accurate alignment of the metering orifices can be attained by drilling and lapping the metering holes of the suspension unit and the fixed part of the valve when they are clamped together. However, for quantity production, it has been found advantageous to machine the two parts separately by using a pair of drilling and reaming jig plates as shown in Fig. 6. This technique insures that the holes match very accurately

at the mating surfaces. Parts thus made are interchangeable, an important factor in quantity production. The accuracy of valves made in this way can be determined best from a displacement-flow curve; a typical one is shown in Fig. 7.

The upward force due to hydraulic pressure acting on the bottom face of the valve plate is the cause of two significant effects: (a) It causes the suspension arms to deflect under compression. This makes metal-to-metal contact between the valve plate and valve body impossible, thereby minimizing valve friction, but it also increases the valve clearance and therefore the quiescent leakage of the valve and the load sensitivity of the actuator. Actual measurements show that the significant factors in suspension-arm deflection are direct compression in the thin sections and column-type bending in the thick sections. The valve was so designed that the total deflection under maximum

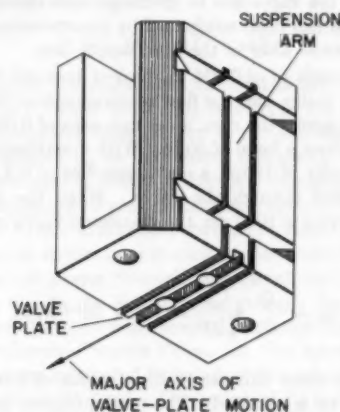


FIG. 5 ISOMETRIC VIEW OF VALVE SUSPENSION UNIT

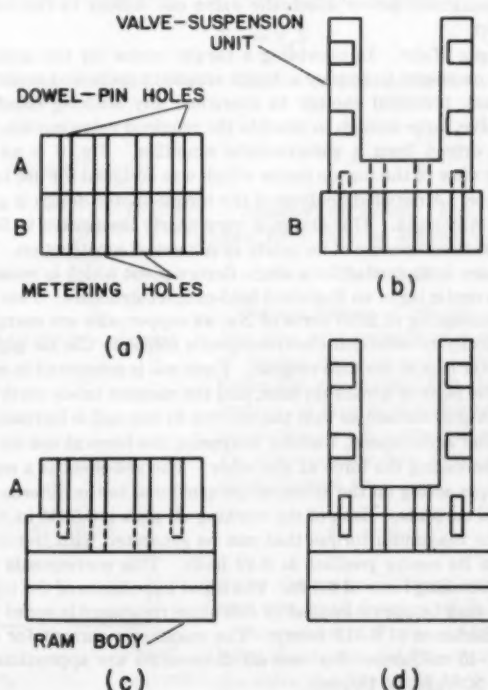


FIG. 6 USE OF DRILLING AND REAMING JIGS IN MAKING SUSPENSION VALVE

pressure force does not exceed 0.0001 in. (b) The upward pressure force can cause the valve to tend to open in one direction or the other, depending on how much the valve plate is deflected relative to its reed pivots, as shown in Fig. 8. When the pressure force is constant, any unbalance due to initial deflection relative to the pivots when the valve is centered can be canceled by an adjustment somewhere in the system or by biasing the torque-motor current, so that the opening effect simply acts like a negative spring of constant stiffness acting on the valve plate. When the pressure varies, as for example, because of supply-pressure variation, the opening force will vary with pressure and the error appearing in ram position will be proportional to the pressure variation. Only proper initial alignment of the centered valve plate relative to its reed pivots will make possible a steady ram position that is insensitive to supply-pressure variations.

Sticking of the valve due to operation with dirt-contaminated hydraulic fluid has been minimized by incorporating a small felt or sintered-bronze filter in the fluid supply line.

A valve motion of ± 0.004 in. proved to work well with the torque-motor design and the flow requirements of the valve. At 2000-psi drop across the ram, a net ram area of 0.025 sq in. is required to produce a force of 50 lb. With a maximum desired no-load ram velocity of 18 ips, a maximum flow of 0.5 cu in. per sec (cis) is required through the valve. When the pressure drop across each orifice is 1000 psi, the maximum valve orifice opening is found from

$$A_o = \frac{Q}{C_d} \sqrt{\frac{\rho}{2P}} = 1.6 \times 10^{-4} \text{ sq in.}$$

In order to produce this area with a stroke of 0.004 in., a port width of 0.04 in. is required. The total quiescent leakage flow to the valve when it is supplied with hydraulic oil at 2000 psi is 0.03 cis, a value which represents a quiescent power drain of 0.009 hp. The maximum power which the valve can deliver to the ram is 0.15 hp.

Torque Motor. In providing a torque motor for the actuator it was necessary to employ a design simple to make and maintain, compact, powerful enough to overcome any sticking effects in the valve, large enough to provide the required valve motion, and easily driven from a vacuum-tube amplifier. Fig. 9 is an isometric view of the torque motor which was designed for use in the actuator. A detailed analysis of the torque-motor design is given in the Appendix. This design is very nearly insensitive to linear accelerations because of its nearly symmetrical construction. The armature is supported by a single flexure pivot which is mounted on the center leg of an E-shaped field-magnet structure. Two coils each consisting of 2050 turns of No. 44 copper wire are energized electrically to establish electromagnetic forces in the air gaps at the outer legs of the field magnet. Each coil is connected in series with the plate of a vacuum tube, and the vacuum tubes are driven in push-pull fashion so that the current in one coil is increased as the other is decreased, thereby increasing the force at one air gap and decreasing the force at the other. The net effect is a couple or torque acting on the armature proportional to the difference of the coil currents. Each of the working air gaps is 0.0045 in. thick and the maximum torque that can be generated with the armature in its center position is 0.19 in.-lb. This corresponds to a valve-stroking force of 0.5 lb. The input impedance of the torque motor may be approximated by 500 ohms resistance in series with an inductance of 0.615 henry. The maximum current for each coil is 40 milliamp. Its over-all dimensions are approximately $\frac{3}{8}$ in. \times $\frac{1}{2}$ in. \times $1\frac{1}{8}$ in.

Ram. The ram consists of a single piece of hardened steel with a piston diameter of $\frac{7}{32}$ in. and a shaft diameter of $\frac{3}{32}$ in. The

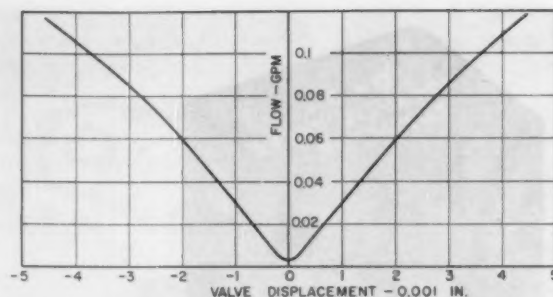


FIG. 7 SUPPLY FLOW VERSUS VALVE DISPLACEMENT FOR SUSPENSION VALVE WITH NO PRESSURE DROP ACROSS THE LOAD. TESTS MADE WITH UNIVIS 40 HYDRAULIC FLUID. PORT WIDTH = 0.040 IN.; SUPPLY PRESSURE = 1750 PSI

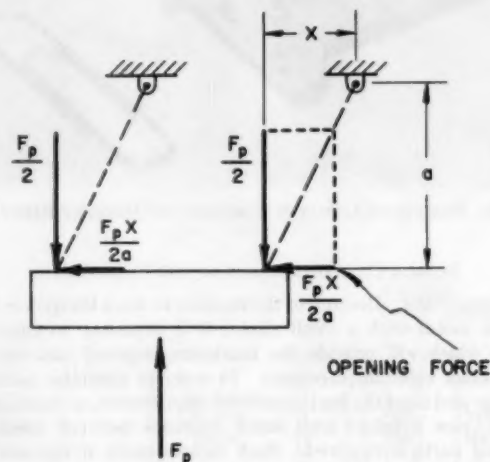


FIG. 8 SKETCH OF VALVE PLATE SHOWING HOW AN OPENING FORCE RESULTS FROM PRESSURE FORCE ACTING ON BOTTOM OF VALVE PLATE

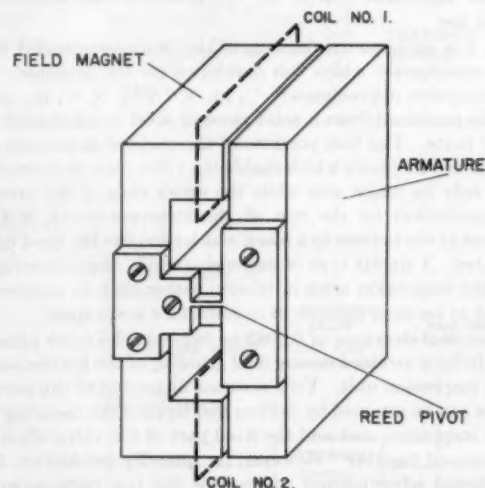


FIG. 9 ISOMETRIC DRAWING OF ELECTROMECHANICAL TORQUE MOTOR

ram shaft is sealed at each end by O-rings as shown in Fig. 10.

Feedback Spring. The feedback mechanism is merely a flat cantilever-type spring rigidly attached to the ram. Since the deflection of this spring is equal to the difference between the ram motion and the valve motion, the force exerted by the spring on the valve plate is proportional to ram motion only when the valve plate has been returned to its center position (as during any steady condition). The effective stiffness of this spring is 31 lb/in.

DYNAMIC ANALYSIS OF ACTUATOR SYSTEM

In order to reveal a number of important facts about the complete system, a dynamic analysis of the actuator system has been made for small changes of all system variables. This analysis,

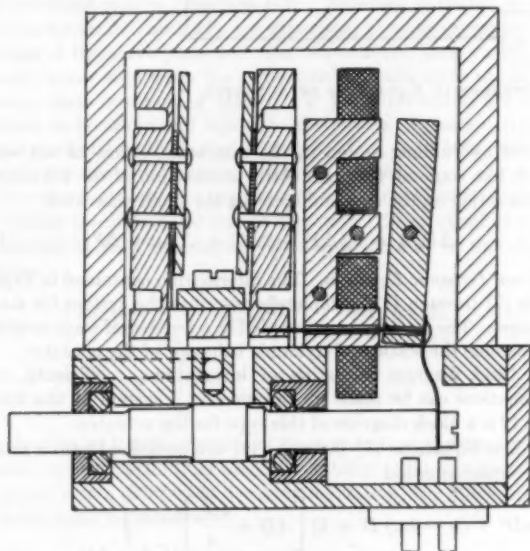


FIG. 10 ASSEMBLY DRAWING OF ELECTROMECHANICAL TORQUE MOTOR

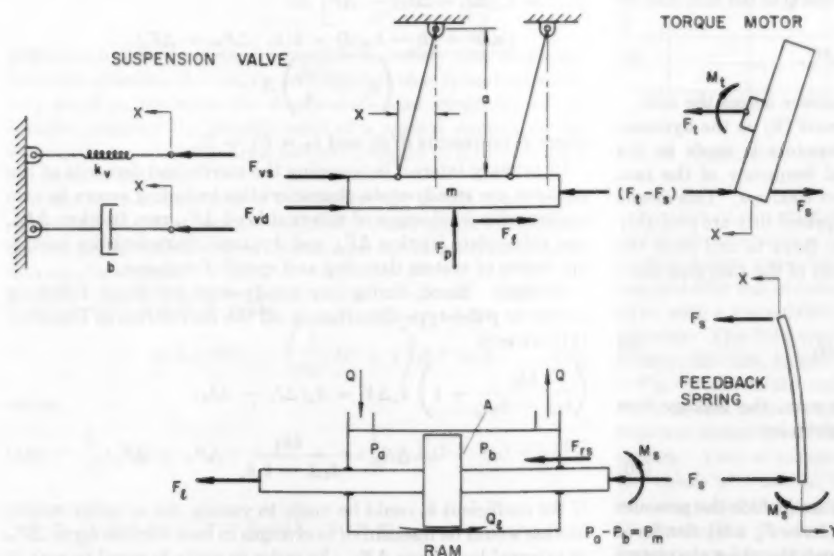


FIG. 11 SKETCH SHOWING FREE-BODY DIAGRAMS OF VARIOUS PARTS OF ACTUATOR SYSTEM

which consists of carefully characterizing each part of the system mathematically and then combining the various component equations into a single mathematical expression, is given in the following paragraphs.

Output Force of Torque Motor. From the detailed analysis in the Appendix, the force developed on the valve plate by the torque motor is found to be a function of both the net difference of the coil currents and of the position of the armature. With the assumption that only small changes occur, the following equation may be used for the torque motor

$$\Delta F_t = k_a(\Delta I_1 - \Delta I_2) - k_b \Delta X \dots \dots \dots [1]$$

where Δ indicates a small change.

Feedback-Spring Force. The feedback spring exerts a force which tends to oppose the force from the torque motor, as shown in Fig. 11. This force is given by

$$\Delta F_s = k_s(\Delta Y + \Delta X) \dots \dots \dots [2]$$

Also acting on the suspension valve plate are (a) an upward pressure force F_p , which is directly related to the system supply pressure; (b) a spring force F_{vs} due to the stiffness of the flexure pivots (represented by an equivalent coil spring of stiffness k); (c) a viscous-damping force F_{vd} due to a damping device which may be incorporated to provide adequate system stability; (d) a friction force F_f due to friction between the valve plate and valve body (usually an erratic type of stick friction which is encountered as dirt-laden oil passes through the valve); (e) a steady-flow force (6) F_{st} due to the fluid flowing through the valve (which tends to move the valve plate toward its center position).

Hydraulic-Supply-Pressure Force. In this valve the pressure force F_p has been estimated to be approximately 20 lb when the supply pressure is 2000 psi. The action of F_p is equivalent to a force equal to $F_p X/a$ acting in the horizontal direction on the valve plate.

Flexure Spring Force. The spring force F_{vs} is given by

$$F_{vs} = k_v X \dots \dots \dots [3]$$

Viscous-Damping Force. The viscous-damping force F_{vd} is given by

$$F_{vd} = b \frac{dX}{dt} \dots \dots \dots [4]$$

Steady-Flow Force. The steady-flow force may be approximated by

$$F_{st} = k_{st} X \dots \dots \dots [5]$$

Friction Force. A friction force F_f also may exist because of the effects of foreign particles getting stuck between the valve plate and the valve body. As long as the pressure drop across the ram is a small fraction of P_s , the unsteady-flow force may be approximated by

$$F_{us} = k_{us} \frac{dX}{dt} \dots \dots \dots [6]$$

when the pressure drop across the ram is a small fraction of P_s .

Summation of Forces. After summing the horizontal forces acting on the valve plate and applying Newton's second law to it, we find in terms of small changes of all variables

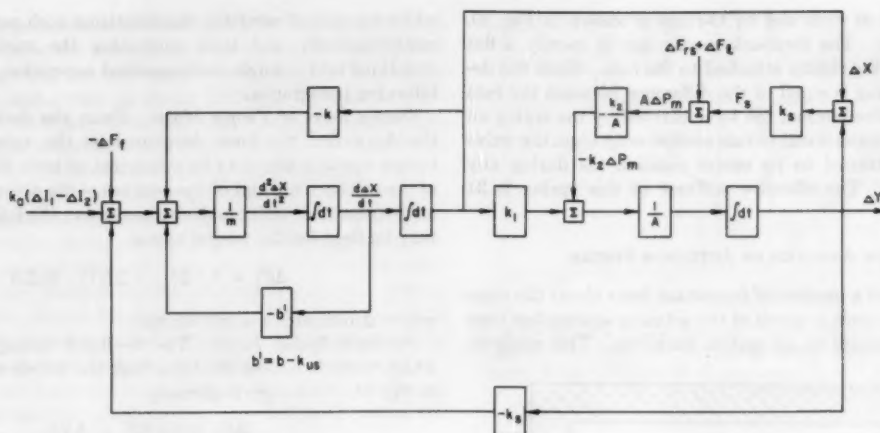


FIG. 12 ANALOG BLOCK DIAGRAM OF BASIC DIFFERENTIAL EQUATIONS OF ACTUATOR

$$k_0(\Delta I_1 - \Delta I_2) - k_b \Delta X - k_s \Delta Y - k_c \Delta X - \frac{F_p \Delta X}{a} - k_c \Delta X - b \frac{d\Delta X}{dt} - \Delta F_f - k_{us} \Delta X + k_{us} \frac{d\Delta X}{dt} = m \frac{d^2 \Delta X}{dt^2}$$

where m is the effective mass of the torque motor and the moving part of the valve.

When terms are collected

$$m \frac{d^2 \Delta X}{dt^2} + (b - k_{us}) \frac{d\Delta X}{dt} + k \Delta X = k_0(\Delta I_1 - \Delta I_2) - k_s \Delta Y - \Delta F_f \dots [7]$$

where

$$k = k_b + k_s + k_c + k_{us} - \frac{F_p}{A}$$

Pressure-Flow Characteristics. The pressure-flow characteristics of the valve are similar to those discussed by Blackburn (8) and Shearer (9). For small changes, the flow rate Q to the ram may be expressed by

$$\Delta Q = k_1 \Delta X - C_1 \Delta P_m \dots [8]$$

where $P_m = P_a - P_b$, the pressure difference across the ram.

An estimate based on the work by Shearer (9) on the dynamic characteristics of valve-controlled servomotors is made in the Appendix and indicates that the natural frequency of the ram system driving a mass of $1/2$ lb is more than 1000 cps. This means that the effects of load mass and fluid compressibility are probably negligible in this system. Therefore the flows to and from the valve are exactly equal to the displacement of the ram plus leakage past the ram; therefore

$$A \frac{d\Delta Y}{dt} = \Delta Q - \Delta Q_l \dots [9]$$

For small pressure difference across the ram, the leakage flow past the ram is viscous (laminar) and is given by

$$\Delta Q_l = C_2 \Delta P_m \dots [10]$$

Ram Forces. The forces acting on the ram include the pressure forces from the hydraulic fluid, the spring force F_s , a friction force F_{rs} , and an external load force F_l . Although the exact characteristics of F_{rs} are not known in detail, this friction force, which is

caused by rubbing action in the ram seals, seems to act very much like coulomb friction. After summation of the horizontal forces acting on the ram and ignoring the small mass load

$$\Delta P_m A = k_s(\Delta Y + \Delta X) + \Delta F_{rs} + \Delta F_l \dots [11]$$

Final Dynamic Equation. The information contained in Equations [7] through [11] completely describes the system for small changes. The set may be assembled in a number of ways in order to examine the static and dynamic behavior of the actuator.

A block diagram consisting of integrations, coefficients, and summations can be made up to represent a system of this kind. Fig. 12 is a block diagram of this type for the actuator.

When Equations [7] through [11] are combined into a single differential equation

$$\left\{ \frac{[mD^2 + (b - k_{us})D + k] \left[AD + \frac{k_2 k_s}{A} \right] + k_2}{k_1 - \frac{k_2 k_s}{A}} \right\} \Delta Y = k_0[\Delta I_1 - \Delta I_2] - \Delta F_f + \frac{[mD^2 + (b - k_{us})D + k] k_2 [\Delta F_{rs} + \Delta F_l]}{\left(k_1 - \frac{k_2 k_s}{A} \right) A} \dots [12]$$

where D represents d/dt , and $k_2 = C_1 + C_2$.

Of primary interest in assessing the merits and demerits of the actuator are steady-state characteristics including errors in ram position due to changes of external load ΔF_l , ram friction ΔF_{rs} , and valve-plate friction ΔF_f , and dynamic characteristics including degree of system damping and speed of response.

Analysis. Since, during any steady-state condition, following a step- or pulse-type disturbance, all the derivatives in Equation [11] are zero

$$\left(\frac{kk_2}{k_1 A - k_2 k_s} + 1 \right) k_s \Delta Y = k_0(\Delta I_1 - \Delta I_2) - \Delta F_f - \frac{k k_2}{k_1 A - k_2 k_s} (\Delta F_{rs} + \Delta F_l) \dots [13]$$

If the coefficient k_2 could be made to vanish, the actuator output motion would be insensitive to changes in ram friction force ΔF_{rs} or external load force ΔF_l . In order to make k_2 equal to zero, it would be necessary to build a valve which is perfectly closed at

its center position and to provide a seal around the ram piston which would allow no leakage past the ram. Because of valve clearance, imperfect location of all of the metering orifices, and rounding of the corners of the metering orifices, a truly closed-center valve cannot be built. The suspension valve employed here nevertheless very nearly approaches being a truly closed-center valve. Leakage past the ram can be reduced to practically zero when a packing is used, but space limitations in this design made it necessary to try to achieve a reasonably good seal around the ram piston by minimizing the clearance between piston and cylinder walls.

The actuator output motion however still remains sensitive to changes in valve friction force ΔF_f even when k_2 is zero, tending to decrease slightly with a small increase of k_2 from zero.

Dynamic operation of the actuator is characterized largely by the left-hand side of Equation [12]. Although variation of the friction forces ΔF_f and ΔF_n are dependent to some extent on the motion of the valve plate and ram, respectively, these variations usually occur only when the respective members come to rest or change their direction of motion. If these variations are considered to be additional inputs to the system, Equation [12] is linear, and the methods and techniques that have been worked out for the study of linear systems can be used.

The characteristic equation for such a linear system is obtained by letting the right-hand side of Equation [12] be equal to zero. When this is done, the characteristic equation may be written as follows

$$\left\{ mD^3 + \left[\frac{k_2 k_s m + (b - k_{us}) A^2}{A^2} \right] D^2 + \left[\frac{k_2 k_s (b - k_{us}) + k A^2}{A^2} \right] D + k_s \left[\frac{k_2 (k - k_s) + k_1 A}{A^2} \right] \right\} \Delta Y = 0 \dots [14]$$

Application of Routh's criterion (10) reveals that in order for the system to have better than marginal stability (that is, respond to an input with a damped or decaying oscillation) the following inequality must be satisfied

$$\left[\frac{k_2 k_s m}{A^2} + (b - k_{us}) \right] \left[\frac{k_2 k_s (b - k_{us})}{A^2} + k \right] > m k_s \left[\frac{k_2 (k - k_s)}{A^2} + \frac{k_1}{A} \right] \dots [15]$$

Although it is theoretically possible to satisfy this inequality when the quantity $(b - k_{us})$ is zero, the fact that k_2 has been made very small to minimize the steady-state load sensitivity of the actuator suggests the possible need of a viscous damper on the valve plate in order to provide sufficient margin to satisfy Routh's criterion for stability.

When the left-hand side of Inequality [15] is just equal to the right-hand side, the system is marginally stable and it will oscillate unendingly in response to an input, and the characteristic equation may be written

$$(a_3 D^3 + a_2 D^2 + a_1 D + a_0) \Delta Y = a_0 (\tau D + 1) \left(\frac{1}{\omega_n^2} D^2 + 1 \right) \Delta Y = 0 \dots [16]$$

where

$$a_0 = \frac{k_s}{A^2} [k_2 (k - k_s) + k_1 A]; \quad a_1 = \frac{1}{A^2} [k_2 k_s (b - k_{us}) + k A^2]$$

$$a_2 = \frac{1}{A^2} [k_2 k_s m + (b - k_{us}) A^2]; \quad a_3 = m$$

and τ and ω_n are the time constant and natural frequency, re-

spectively, of the system. The values of τ and ω_n are useful in making an estimate of the system's speed of response. By equating the coefficients of terms with like powers of D in Equation [16], the expressions for τ and ω_n are found to be

$$\tau = \frac{k_2 k_s (b - k_{us}) + k A^2}{[k_2 (k - k_s) + k_1 A] k_s} \dots [17]$$

$$\omega_n = \sqrt{\frac{[k_2 (k - k_s) + k_1 A] k_s}{k_2 k_s m + (b - k_{us}) A^2}} \dots [18]$$

Because it is usually desirable to have a system which responds with a rapidly decaying oscillation rather than with an unending oscillation, the time constant and natural frequency of such a system will differ somewhat from τ and ω_n given in Equations [17] and [18], but τ and ω_n are useful if only a rough estimate of speed of response is required.

EXPERIMENTAL WORK

Dynamic Characteristics. When the actuator was first assembled, the viscous damper on the valve plate was omitted, and the system was found to oscillate with increasing amplitude when subjected to a small disturbance at any point in the system. In order to provide adequate system stability, viscous damping was employed on the valve plate. This damping action was accomplished by attaching a small plate to each of the vertical suspension arms in such a way that it covered, but did not touch, the adjacent solid vertical supports as shown in Fig. 10. The space between the damping plates and vertical supports is filled with oil which exerts a damping force tending to oppose valve-plate motion as the thickness of the space is changed by valve-plate motion.

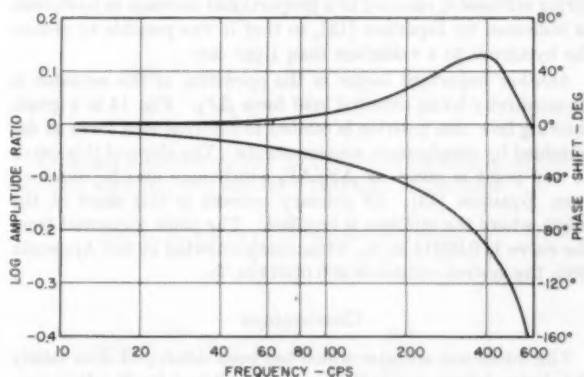


FIG. 13 AMPLITUDE AND PHASE-SHIFT CHARACTERISTICS OF MINIATURE ACTUATOR

Because of the experimental difficulties anticipated in making useful measurements of the characteristics of this damping device, its final design was achieved by trial and error. Satisfactory system stability was obtained with a total effective plate area of 0.25 sq in. with a space 0.004 in. thick with the valve plate in its center position. The fluid used in the dampers is the same as that used to drive the ram, namely, Univis 40.

Fig. 13 shows the amplitude and phase-lag characteristics of the complete actuator system that were obtained from frequency-response measurements after the viscous-damping plates were installed. A set of calculations, given in the Appendix, shows that the values of τ and ω_n for this system are 0.0004 sec and 3500 rad/sec, respectively, when the system is marginally stable. Fig. 13 shows that the real (well-damped) system has a somewhat lower natural frequency than the analysis of a marginally stable system indicates. The graphed frequency-response characteris-

tics include the dynamic characteristics of the electronic amplifier and torque motor. Frequency-response measurements relating torque-motor current to amplifier input signal revealed very little lag or attenuation up to 400 cps.

Since the speed of response of the actuator is more than adequate to meet the requirements, it was not necessary to seek an optimum system speed of response. A thorough study of the fundamental equations together with an analog study based on the block diagram of Fig. 12 would be useful to the designer who wished to arrive at a design with optimum speed of response.

Static Characteristics. The relationship between actuator output position and input current during very low-frequency variations of the input is closely linked to the steady-state characteristics discussed in the analysis section. All of the derivatives in Equation [12] are small in this case, and the steady-state Equation [13] may be employed to gain an insight into the factors which cause a hysteresis effect to exist between the input and the output of the actuator. The friction forces ΔF_f and ΔF_n may be thought of as consisting of coulomb friction. As such, they are nearly independent of velocity of relative motion and always act to oppose motion so that they change rapidly from a negative value to a positive value as relative velocity changes from a positive value to a negative value. Thus, during a test at very low frequency, these forces change only at the extreme positions of the valve and ram, respectively. Because of the very low force levels involved, these forces have not been measured directly. The experimental data that have been derived from low-frequency response tests, however, do indicate that this concept of coulomb friction together with Equation [13] does describe the important effects which take place. As an example, decreasing the net spring stiffness k of the torque motor and valve relative to the feedback-spring stiffness k_f , resulted in a proportional decrease in hysteresis as indicated by Equation [13], so that it was possible to reduce the hysteresis to a value less than 1 per cent.

Another important factor in the operation of the actuator is its sensitivity to an external load force ΔF_l . Fig. 14 is a graph showing how ram position is related to external load force as determined by steady-state measurements. The slope of this curve at any point is given by $\Delta Y/\Delta F_l$ which may also be obtained from Equation [13]. Of primary concern is this slope at the origin where the stiffness is smallest. The slope measured from the curve is 0.00018 in/lb, while that computed in the Appendix from the system constants is 0.00019 in/lb.

CONCLUSIONS

The miniature actuator which has been developed does satisfy the design requirements that were established for it. It is considerably smaller in size than the smallest torque motor which would be required to provide the same output, and its response is considerably faster than such a torque motor would have. As predicted by a dynamic analysis of the system, a viscous damper on the valve plate may be employed to provide adequate system stability.

Extensive tests that have been conducted by operating the actuator with oil contaminated by dirt and miscellaneous foreign particles reveal that the control valve is sensitive to highly contaminated hydraulic fluid, but that small felt or sintered-bronze filters serve as adequate means of preventing valve sticking when the fluid is highly contaminated. Since the maximum flow to the valve is very small, the filter is accordingly small and may be incorporated in the actuator without appreciably increasing its size. The operation to date has shown that the actuator with a suitable filter never fails to operate because of dirty hydraulic fluid.

Preliminary tests with the actuator operating on compressed air at 1000 psi instead of hydraulic oil indicate that development

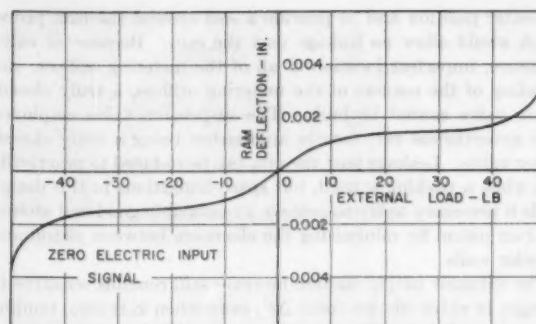


FIG. 14 GRAPH SHOWING EXPERIMENTAL MEASUREMENTS OF RAM DEFLECTION VERSUS EXTERNAL LOAD

of an electropneumatic actuator with speed of response nearly as fast as the electrohydraulic actuator is feasible.

ACKNOWLEDGMENT

The authors wish to thank the staff of the Dynamic Analysis and Control Laboratory for their encouragement and support of the development work that is reported here. This work was supported by funds from the U. S. Navy Bureau of Ordnance Contract NOrd 9661 with the Division of Industrial Cooperation of the Massachusetts Institute of Technology.

Special thanks are due Prof. R. H. Frazier of the Department of Electrical Engineering for his assistance on torque-motor design and to Mr. C. W. Gould for his thorough work in the laboratory.

BIBLIOGRAPHY

- 1 "Design and Operating Instructions for the DACL Servo-Valve Models MX and XA," by R. D. Atchley and R. W. Bond, Dynamic Analysis and Control Laboratory Report No. 50, the Massachusetts Institute of Technology, Cambridge, Mass., September, 1949.
- 2 "An Electric Valve Actuator for Hydraulic Servomechanisms," by R. H. Frazier and R. D. Atchley, Meteor Report No. 42, the Massachusetts Institute of Technology, Cambridge, Mass., September, 1949.
- 3 "A Permanent-Magnet-Type Electric Actuator for Servovalves," by R. H. Frazier and R. D. Atchley, Dynamic Analysis and Control Laboratory Report No. 66, the Massachusetts Institute of Technology, Cambridge, Mass., June, 1952.
- 4 "A Study of Permanent-Magnet Torque Motors," by J. F. Dunn, Jr., Dynamic Analysis and Control Laboratory Research Memorandum No. R.M. 6387-5, the Massachusetts Institute of Technology, Cambridge, Mass., April, 1954.
- 5 "Contributions to Hydraulic Control—6, New Valve Configurations for High-Performance Hydraulic and Pneumatic Systems," by S.-Y. Lee, Trans. ASME, vol. 76, 1954, pp. 905-911.
- 6 "Contributions to Hydraulic Control—1, Steady-State Axial Forces on Control-Valve Pistons," by S.-Y. Lee and J. F. Blackburn, Trans. ASME, vol. 74, 1952, pp. 1005-1016.
- 7 "Contributions to Hydraulic Control—2, Transient-Flow Forces and Valve Instability," by S.-Y. Lee and J. F. Blackburn, Trans. ASME, vol. 74, 1952, pp. 1013-1016.
- 8 "Contributions to Hydraulic Control—3, Pressure-Flow Relationships for 4-Way Valves," by J. F. Blackburn, Trans. ASME, vol. 75, 1953, pp. 1163-1170.
- 9 "Dynamic Characteristics of Valve-Controlled Hydraulic Servomotors," by J. L. Shearer, Trans. ASME, vol. 76, 1954, pp. 895-903.
- 10 "Mechanical Vibrations," by J. P. Den Hartog, McGraw-Hill Book Company, Inc., New York, N. Y., third edition, 1947, pp. 350-354.

Appendix

TORQUE-MOTOR ANALYSIS

Fig. 15 is a drawing of the torque-motor configuration em-

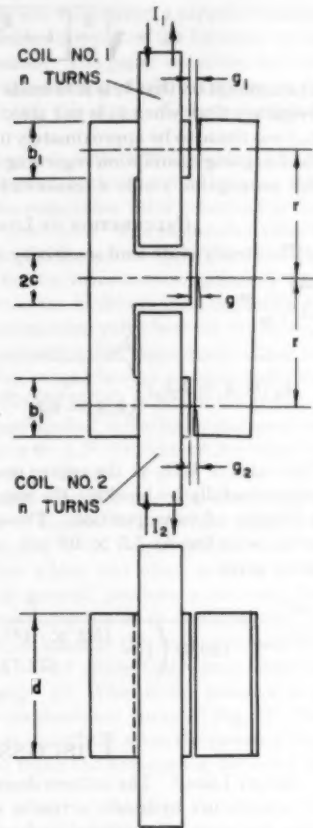


FIG. 15 DRAWING OF TORQUE MOTOR

played in the actuator. The electromagnetic force exerted on the armature at each of the active gaps is given by

$$F_1 = \frac{k_1}{b_1 d} (\phi_1)^2 \quad [19]$$

$$F_2 = \frac{k_2}{b_1 d} (\phi_2)^2 \quad [20]$$

In order to consider small changes from a set of initial conditions, Equations [19] and [20] may be written

$$F_{1i} + \Delta F_1 = \frac{k_1}{b_1 d} (\phi_{1i} + \Delta \phi_1)^2 \quad [21]$$

$$F_{2i} + \Delta F_2 = \frac{k_2}{b_1 d} (\phi_{2i} + \Delta \phi_2)^2 \quad [22]$$

so that when Equation [19] is combined with [21], and [20] is combined with [22]

$$\Delta F_1 = \frac{k_1}{b_1 d} [2\phi_{1i}\Delta\phi_1 + (\Delta\phi_1)^2] \quad [23]$$

$$\Delta F_2 = \frac{k_2}{b_1 d} [2\phi_{2i}\Delta\phi_2 + (\Delta\phi_2)^2] \quad [24]$$

When $\Delta\phi_1$ and $\Delta\phi_2$ are only a few per cent of ϕ_{1i} and ϕ_{2i} , the terms $(\Delta\phi_1)^2$ and $(\Delta\phi_2)^2$ may be neglected

$$\Delta F_1 = \frac{2k_1}{b_1 d} \phi_{1i}\Delta\phi_1 \quad [25]$$

$$\Delta F_2 = \frac{2k_2}{b_1 d} \phi_{2i}\Delta\phi_2 \quad [26]$$

The flux in each air gap is a function of the magnetomotive force generated by the current flowing in each coil and the total reluctance of each magnetic circuit. As a first approximation the reluctance of the magnetic material in the field magnet and armature may be neglected so that the reluctance of each magnetic circuit is due to its air gaps. Half the area of the air gap at the center leg may be considered to contain ϕ_1 and the other half to contain ϕ_2 . The following equations then may be used

$$\phi_1 = k_3 \frac{nI_1}{\frac{g_1}{b_1 d} + \frac{g}{cd}} \quad [27]$$

$$\phi_2 = k_3 \frac{nI_2}{\frac{g_2}{b_1 d} + \frac{g}{cd}} \quad [28]$$

The values of g_1 and g_2 are the same and equal to g_0 when the armature is centered. When the armature is away from center

$$g_1 = g_0 + X \quad [29]$$

$$g_2 = g_0 - X \quad [30]$$

Combining Equations [27] with [29] and Equation [28] with [30] and using logarithmic differentials

$$\Delta\phi_1 = \phi_{1i} \left(\frac{\Delta I_1}{I_{1i}} - \frac{\Delta X}{g_0 + X_i + \frac{b_1}{c} g} \right) \quad [31]$$

$$\Delta\phi_2 = \phi_{2i} \left(\frac{\Delta I_2}{I_{2i}} + \frac{\Delta X}{g_0 - X_i + \frac{b_1}{c} g} \right) \quad [32]$$

The changes in ϕ_1 and ϕ_2 may be eliminated by combining Equation [25] with [31] and Equation [26] with [32] to give

$$\Delta F_1 = \frac{2k_1}{b_1 d} (\phi_{1i})^2 \left(\frac{\Delta I_1}{I_{1i}} - \frac{\Delta X}{g_0 + X_i + \frac{b_1}{c} g} \right) \quad [33]$$

$$\Delta F_2 = \frac{2k_2}{b_1 d} (\phi_{2i})^2 \left(\frac{\Delta I_2}{I_{2i}} + \frac{\Delta X}{g_0 - X_i + \frac{b_1}{c} g} \right) \quad [34]$$

Summing the forces acting on the torque-motor armature

$$\Delta F_t = \Delta F_2 - \Delta F_1 - k_{rp}\Delta X \quad [35]$$

When the initial conditions are $I_{1i} = I_{2i} = I_i$, $X_i = 0$, and $F_{ti} = 0$

$$\Delta F_t = k_a(\Delta I_1 - \Delta I_2) - k_b\Delta X \quad [36]$$

where

$$k_a = \frac{2k_1 k_3^2 n^2 I_i b_1 d}{\left(g_0 + \frac{b_1}{c} g \right)^2} \quad [37]$$

$$k_b = k_{rp} - \frac{4k_1 k_3^2 n^2 I_i^2 b_1 d}{\left(g_0 + \frac{b_1}{c} g \right)^2} \quad [38]$$

For the miniature actuator, k_a may be calculated using

$$k_a k_b^2 = 1.4 \times 10^{-7} \text{ lb/(amp turn)}^2$$

and the values of the constants given in the list of nomenclature

$$k_a = \frac{2(1.4 \times 10^{-7})(2000)^2(0.02)(0.125)(0.25)}{0.0045 + \frac{(0.125)(0.0015)^2}{(0.063)}} = 11 \text{ lb/amp} \quad [39]$$

and k_b may be calculated

$$k_b = (43.0) - \frac{4(1.4 \times 10^{-7})(2000)^2(0.02)(0.125)(0.25)}{(0.0075)^2} = 43.0 - 63.7 = -20.7 \text{ lb/in.} \quad [40]$$

The maximum force developed by the torque motor in its center position may be determined by using Equations [19] and [27] because the current in one coil is a maximum of 40 ma and the current in the other coil is zero

$$F_t = \frac{k_a k_b^2 n^2 I_1^2 b_1 d}{\left(g_0 + \frac{b_1 g}{c}\right)^2} \quad [41]$$

or

$$F_t = \frac{(1.4 \times 10^{-7})(2000)^2(0.04)^2(0.125)(0.25)}{(0.0075)^2} = 0.5 \text{ lb.} \quad [42]$$

NATURAL FREQUENCY OF RAM AND MASS LOAD

With the constant k_2 very small, the natural frequency of the ram and mass load may be approximated as follows

$$\omega_{ns} = \sqrt{\frac{2A^2 \beta}{V m_1}} \quad [43]$$

Calculating ω_{ns} when m_1 weighs 0.5 lb

$$\omega_{ns} = \sqrt{\frac{(2)(0.025)^2(2.5 \times 10^6)(386)}{(0.0025)(0.5)}} = 8870 \text{ rad/sec.} \quad [44]$$

CALCULATION OF τ AND ω_n

For unending oscillation, Routh's criterion gives

$$[k_2 k_a m + (b - k_{as})A^2][k_2 k_b(b - k_{as}) + kA^2] = m k_a A^2 [k_2(k - k_a) + k_1 A] \quad [45]$$

Assuming k_2 is zero

$$(b - k_{as}) = \frac{k_1 k_a m}{k A} \quad [46]$$

and

$$\tau = \frac{kA}{k_1 k_2} \quad [47]$$

Solving for τ

$$\tau = \frac{(73.7)(0.025)}{(168)(29.6)} = 0.0004 \text{ sec.} \quad [48]$$

Continuing with the assumption that $k_2 = 0$

$$\omega_n = \sqrt{\frac{k}{m}} = \sqrt{\frac{(73.7)}{(6 \times 10^6)}} = 3500 \text{ rad/sec.} \quad [49]$$

A more exact expression for ω_n is given by Equation [18] or by its counterpart

$$\omega_n = \sqrt{\frac{k}{m} \left[1 + \frac{k_a k_2 (b - k_{as})}{kA^2} \right]}$$

The assumption that k_2 is zero leads to a somewhat higher natural frequency than when k_2 is not zero, but the term $(k_a k_2)/kA^2(b - k_{as})$ was found to be approximately 0.0007 in this case. Therefore the foregoing assumption regarding k_2 is justified. In any event, this assumption yields a conservative value of ω_n .

CALCULATION OF LOAD SENSITIVITY

The steady-state load sensitivity is given by

$$\frac{\Delta Y}{\Delta F_t} = \frac{k k_2}{k_1 (k_1 A - k_2 k_a) \left(\frac{k k_2}{k_1 A - k_2 k_a} + 1 \right)} = \frac{1}{k_a \left(1 + \frac{k_1 A}{k_2 k} - \frac{k_2}{k} \right)} \quad [50]$$

The value of k_1/k_2 at the center position of the valve was found experimentally by blocking the ram and finding load pressure as a function of valve position. These measurements showed k_1/k_2 could be as low as 5.3×10^6 pci. Calculation of the load sensitivity gives

$$\frac{\Delta Y}{\Delta F_t} = \frac{1}{(29.6) \left(1 + \frac{(5.3 \times 10^6)(0.025)}{(73.7)} - \frac{(29.6)}{(73.7)} \right)} = 1.88 \times 10^{-4} \text{ in/lb.} \quad [51]$$

Discussion

SIDNEY LEES.⁴ The authors describe the successful completion of a miniature hydraulic actuator using flat valve plates rather than the more conventional hydraulic valve design. They have paid careful attention to detail with respect to both structure and performance. One cannot take issue with the design for which the authors merit congratulations. The writer finds a few points of difference with them in the presentation. For example, a device which generates force with respect to an input current is not one the writer would characterize as a torque motor. It would seem that a name such as force motor is more suitable for this particular application.

It is estimated that the size of this device is a cube $1\frac{1}{2}$ in. on edge. To explain its performance the authors found it necessary to employ 68 symbols conveniently collected at the beginning of the paper. The symbols are all necessary; however, for lack of a well-developed self-defining notation, it is found that the symbol b is used to represent the viscous damping coefficient while b_1 is the height of outer legs of the torque-motor field magnet. A profusion of k 's with numerical subscripts refer to valve coefficient, fluid-compliance coefficient, electromagnetic constant, torque-motor coefficient, etc. The same spatial displacement direction is referred to as X or Y depending on whether it is the valve or the ram which is involved. The writer found it necessary continually to refer to the nomenclature in order to be able to read the paper. Furthermore, he found it rather difficult to keep the ideas straight, since he was never sure the symbol meant what he thought it did. A partial use of self-defining symbolism is exhibited but the authors stopped well short of completion. It is believed the paper is a good illustration of where self-defining notation is very applicable.

R. E. MICHEL.⁵ The authors have presented an interesting

⁴ Massachusetts Institute of Technology, Cambridge, Mass.

⁵ Research Project Engineer, Minneapolis-Honeywell Regulator Company, Minneapolis, Minn.

and unique method of valving and have given a capable development of the analytical expressions describing the behavior of the complete hydraulic servoactuator. The paper is concise and well organized and the information presented should prove to be very useful to the designer of this type of electrohydraulic actuator.

The designer who is concerned with valve sticking due to operation with dirt-contaminated hydraulic fluid should consider the possible merits of the suspension valve described in this paper over the conventional spool type. The suspension valve is no less susceptible to silting than the piston-type valve but the phenomena of lateral forces due to valve-land irregularities that are inherent or caused by dirt in the hydraulic fluid cannot result in high friction loads on the suspension valve because the moving element is restrained from contacting the valve body. It is interesting to note that as silting takes place on a piston valve the configuration is effectively changed to one in which the piston is tapered with the higher pressure applied to the larger end resulting in a decentering force. A paper by J. F. Blackburn⁶ describes the phenomena of these side forces in detail.

A few minor questions arise: (1) What is the amplitude of ram motion of Fig. 13? It is of interest to know the amplitude of ram displacement corresponding to the "log-amplitude ratio" of zero. Inherent nonlinearities which one often neglects in an analysis of a system will, in general, produce a resonant frequency which is lower at larger amplitudes of displacement. It would be of further interest to know if the frequency-response characteristic of this system varied appreciably when different amplitudes of motion are used. (2) What is the pressure drop across the valve in the flow-displacement curve of Fig. 7? The curve would seemingly be most significant when the pressure drop across the valve is such as to make the horsepower delivered by the ram at a maximum.

D. G. O'BRIEN.⁷ The paper is well written; the development is logical and easily followed. The authors are to be congratulated for their contribution toward the development of miniaturized components for the hydraulic-control field. A great deal of effort has, of course, already been devoted to the problems of miniaturizing electronic circuitry but, because hydraulic mechanisms are, in general, quite small for their power capabilities when compared with other systems which perform similar functions, attempts to miniaturize hydraulic systems have not been widely prevalent up to the present time.

Since many potential applications for a device of the type described in the paper have no significant limitations as far as allowable quiescent internal oil leakage is concerned, the thought has occurred to the writer that perhaps the internal leakage might be increased somewhat in such a way as to maintain at the existing level most performance specifications, while at the same time allowing some production economies to be made which may significantly lower the cost of the completed device. Specifically, it is suggested that:

1 The valve's coefficient K_1 may be increased somewhat without too much deterioration in the good linearity shown in Fig. 7.

2 The area of the actuating piston as well as the leakage coefficients C_1 and C_2 may also be increased.

3 If the proper choice of the foregoing constants is made, it should be possible to maintain at very nearly the present level the steady-state load sensitivity as given in Equation [50], while in-

creasing the actuator's damping to the point where possibly the viscous damper may be eliminated.

4 The possible elimination of the damper as well as the increased clearance between mating parts that may be allowed in view of the increased leakage should make the device less expensive to fabricate.

It is realized by the writer that suggestions for possible improvement of the device without access to much more information (particularly experimental) than was presented in the paper may be out of order. The writer's experience to date in the field of hydraulic controls, however, has indicated a continuing need for hydraulic servocomponents which are less expensive than those now commercially available, and the fact that in many applications the excellent dynamic capabilities of available components are not needed or, in many cases, wanted.

The information which is available to the authors of the paper, as a result of their work on this device, should make it possible for them rather easily to determine whether or not the changes proposed by the present writer are feasible.

F. B. TUTEUR.⁸ The clear and logical description and analysis of this new hydraulic actuator is certainly a most worthwhile contribution to the field of hydraulic control. As anyone who has had any experience in the design of hydraulic valves knows, one of the most difficult problems is the construction of a high-pressure valve which is capable of being operated by an electric solenoid of only limited force capabilities. In the ordinary piston valves one always encounters lateral forces which tend to push the valve piston against one side of the sleeve and thus produce very large coulomb friction forces.⁹ In addition, valves are usually subject to hydraulic reaction forces tending to close the metering orifices unless special precautions,¹⁰ that are never completely effective, are taken against them. Apparently the authors have in good part overcome both difficulties by their ingenious sliding valve design. The pressure under the valve forces the sliding plate off the main body by a small, but precisely controllable amount, thus reducing friction to a minimum without introducing excessive leakage; at the same time, the off-center force resulting from the suspension tends to open the valve, thus canceling, in part, the troublesome hydraulic reaction forces. A further welcome feature is the relative ease of constructing the metering orifice with great accuracy; this is usually also a difficult matter in piston-valve design.

It would seem that the steady-flow force F_{st} should be a function not only of valve displacement but also of flow through the valve,^{9,10} thus a term proportional to dY/dt would enter the equations. Possibly this term is negligibly small in the system considered. In any case, it would seem to be possible by proper choice of the distance a and the spring constant k_s and k_v to make the effective valve spring constant k equal to zero. Under these conditions the loop would contain an additional integration so that the steady-state load sensitivity $\Delta Y/\Delta F_{l/ss}$ could be reduced to zero. This fact is also indicated by Equation [50]. The stability condition, Equation [15], would not be changed very vitally by letting $k = 0$, so that the system could still be stable. Thus it seems that a superior system might result.

The actual curve of steady-state load sensitivity, Fig. 14, indicates some definite nonlinear effects, not completely explained by the linear analysis given. Apparently what takes place is that near-zero deflection, the valve constant k_1 , is reduced (as shown in Fig. 7) and also the leakage constant k_2 may be increased. This, from

⁶ "Contributions to Hydraulic Control—5, Lateral Forces on Hydraulic Pistons," by J. F. Blackburn, Trans. ASME, vol. 75, 1953, pp. 1175-1180.

⁷ Vice-President, Engineering & Production, Midwestern Instruments, Tulsa, Okla.

⁸ Department of Electrical Engineering, Yale University, New Haven, Conn.

⁹ Reference (6) of paper.

¹⁰ Reference (7) of paper.

Equation [50], would indicate the larger slope of $\Delta Y/\Delta F_{1/2}$ near-zero deflection as is borne out by Fig. 14.

Interesting additional data would be step-function response to input and to load disturbances.

AUTHORS' CLOSURE

The comments of Professor Lees regarding usage of the term torque motor are well taken but since the term seems to be so firmly entrenched in the minds of the people working in this field, it is not likely that it will be changed.

In response to Mr. Michel's questions: The amplitude of the ram motion at low-frequency operation for the curves given in Fig. 13 was about 0.008 in. peak-to-peak. The response characteristics did not change appreciably with amplitude, but at very large amplitudes (larger than 0.020 in. peak-to-peak) the effects of velocity limiting due to saturation of the control valve were evident. The caption for Fig. 7 states that the supply pressure was 1750 psi. Because there was no pressure drop across the load, the drop across the valve was 1750 psi. Since this curve was obtained in order to determine linearity of the metering orifices and flow passages as well as to determine the quiescent leakage flow through the valve, there was no reason to maximize load power output.

In answer to Mr. O'Brien's queries, the authors agree that there are many applications in which less precision in the manufacture of the valve and other components is required. Each application should be considered on its own merits. It is hoped that the attempt to define valve performance with simple coefficients such as k_1 , c_1 , and c_2 will assist competent system designers in selecting

the proper type and quality of components for each application. The need for viscous damping at the valve seems to arise whenever the speed of response of the entire system approaches the natural frequency of the valve and torque-motor assembly. In systems where a relatively low speed of response is required, the damper can be eliminated with considerable simplification of the design. This does not seem to be a case in which the allowance of greater valve leakage will provide the type of damping required. This can be readily seen in Fig. 12 which shows clearly the greatly different roles played by valve damping b and valve and ram leakage k_2 .

As Professor Tuteur suggests, the steady-flow force is a function not only of the valve opening but also of the pressure drop across the valve. In cases where the pressure drop across the ram is appreciable, it would seem advisable to take this effect into account in the proper manner. It is possible to consider the steady-flow force as a function of valve opening and ram flow, but when this is done it is not so easy to visualize the effects of ram load conditions.

With reference to making k equal to zero, the authors do not recommend trying to do this in a system of this kind because of resulting extra integration which makes it very difficult to attain adequate system stability without increasing valve and ram leakage.

Detailed pressure-stroke characteristics of the valve, mentioned only briefly in the paper, do show that definite nonlinear effects exist when large changes in ram pressure drop occur. This topic could well be discussed at length in a future paper.

Residual Grinding Stresses in Hardened Steel¹

By H. R. LETNER,² PITTSBURGH, PA.

Residual stresses resulting from surface grinding a hardened ball-bearing-type steel under closely controlled conditions were measured by the deflection method and the effects of wheel grade, unit downfeed, and grinding fluid upon the stresses generated were studied. The stress gradients close to the surface were found to be quite steep and, in a number of instances, both components of the biaxial stress were compressive at the surface. This behavior contrasts with previous dry-grinding experiments on annealed steel and indicates the possibility of generating compressive surface stresses in hardened steel by grinding.

INTRODUCTION

PUBLISHED information concerning the magnitude and distribution of residual grinding stresses in hardened steel is very meager. Glickman (1)³ and his associates made measurements on hardened high-chromium stainless steel but made no attempt to correlate the stresses with grinding conditions. Since the severity of the grinding operation very likely has a great deal to do with the magnitude, and perhaps the distribution, of the induced stresses, there is a need for experiments in which the grinding conditions are closely controlled. Some information of this kind has been published on grinding stresses in annealed steel; it is known, for example, that the stresses are at least biaxial and that their magnitude and depth of penetration depend upon the wheel depth of cut (2, 3).

In the absence of specific information concerning the nature of grinding stresses in hardened steel, it is frequently assumed that they have high tensile values and that, because of this, the service life of ground parts, under conditions of fluctuating applied stress, is impaired by the grinding operation. A comprehensive study has been started to determine just how grinding stresses and fatigue properties are actually affected by grinding conditions. The present paper is concerned with a preliminary study of the effect of several combinations of grinding conditions, differing in severity, upon residual stresses induced in hardened steel.

SPECIMENS

The steel was a ball-bearing type, similar to AISI 52100, having the following percentages of alloying elements: 1.01 carbon, 1.25 chromium, 0.27 manganese, 0.19 vanadium, 0.36 silicon, 0.011 sulphur, and 0.011 phosphorus. It was furnished by Vanadium Alloys Steel Company as a special heat to duplicate as closely as

¹ A contribution from the L. Leslie Byers Memorial Fellowship sustained at Mellon Institute by the Grinding Wheel Institute and the Abrasive Grain Association.

² Senior Fellow, Mellon Institute.

³ Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 2, 1954. Paper No. 54-A-56.

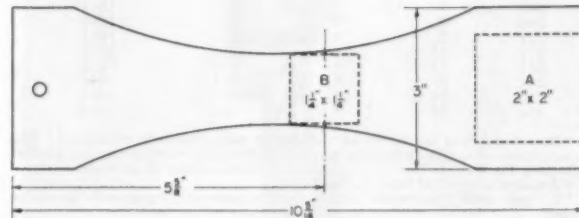


FIG. 1 DIAGRAM OF TEST BAR SHOWING POSITIONS FROM WHICH SPECIMENS FOR STRESS MEASUREMENTS WERE CUT

possible the steel used earlier in the fatigue studies reported by Tarasov and Grover (4). Flat fatigue bars of the shape shown by the solid line in Fig. 1 were prepared from 0.200-in.-thick hot-rolled and annealed sheet, their long dimension being taken in the rolling direction. After equal amounts had been ground off both sides to remove decarburization completely, the bars were austenitized 10 min at 1550 F in a salt bath, quenched in oil, and tempered 2 hr at 450 F to a Rockwell hardness of C59. A check revealed that at a depth of 0.020 in., the thickness subsequently to be removed from each side by grinding, the residual stresses resulting from heat-treatment were too small to cause any appreciable error in determining the grinding stresses.

Approximately 0.010 in. was ground carefully from each surface of the bars, under machine-shop conditions, to bring them to a uniform thickness of 0.140 in. The grinding was done on a horizontal spindle, reciprocating-table type of surface grinder equipped with a 10-in. \times 3-in. \times 3-in. 46-grit aluminum-oxide abrasive wheel, soft grade, open structure, and vitrified bond (32A46-F12VBEP). A grinding fluid consisting of a 2 per cent concentration of a general-purpose miscible oil in water was used. The residual stresses in the surfaces of a typical bar were checked at this stage to be sure that the stresses induced in grinding the bars to the 0.140 in. thickness would not influence the stresses induced by the test grinding to follow.

GRINDING TESTS

The grinding tests were designed to give information on the effect of wheel grade, downfeed, and grinding fluid upon residual stresses. They were conducted on a horizontal spindle, reciprocating-table type of surface grinder under laboratory conditions. Grinding was done on the periphery of 8-in. \times 0.75-in. \times 1.25-in. medium-structure wheels of letter grades H to M,⁴ containing 46-grit white fused aluminum-oxide abrasive in a vitrified bond (38A46-H8VBE to 38A46-M8VBE, inclusive). The wheel speed was 6000 sfpm, traverse speed 60 fpm, and the unit cross-feed 0.050 in. per table stroke.

The bars were ground lengthwise, one at a time, the top surface (identifying numbers on edges erect) first and then the bottom. Approximately 0.010 in. was removed from each surface to give a final thickness of 0.120 in. The wheel was freshly dressed before grinding each surface, using a 0.001-in. downfeed and a moder-

⁴ Letter grades extend from A to Z in order of increasing hardness. For the surface grinding of a hardened steel, H and M-grade wheels would be considered soft and very hard, respectively.

TABLE 1 SUMMARY OF CONDITIONS STUDIED IN GRINDING TESTS

Bar no.	Wheel grade	Grinding fluid ^c	Downfeeds	
			Top	Bottom
203.....	H	e	4-1	4-1
			8-0.5	8-0.5
			6-0.25	6-0.25
			6-0.1	5-0.1
204.....	H	e	4-1	4-1
			12-0.5	12-0.5
205.....	H	e	12-1	12-1
206.....	H	e	6-2	6-2
207.....	I	e	12-1	12-1
209.....	J	e	12-1	11-1
210.....	K	e	12-1	12-1
211.....	L	e	12-1	10-1
212.....	M	e	12-1	11-1
213.....	M	e	6-2	5-2
214.....	M	d	12-1	8-1
215.....	M	e	10-1	10-1
216.....	M	e	5-2	5-2

^a The wheel was fed down at the end of each complete cross-feed. The number preceding the dash is the number of complete cross-feeds; the number following the dash is the unit downfeed in thousandths of an inch.

^b Accidental error of feed mechanism.

^c 2.5 per cent transparent emulsion of a general-purpose soluble oil in water.

^d Air.

^e Highly activated sulphochlorinated grinding oil, 80 sec Saybolt Universal Viscosity at 100 F.

ately rapid cross-feed of the diamond to leave the wheel face open. The wheel condition at any stage of grinding was thus the same for both surfaces.

The grinding conditions which were studied are summarized in Table 1. The conditions for many of the tests were more severe than would be encountered in good grinding practice because of the excessive hardness of some of the wheels. In a comparable commercial operation, wheels equivalent to the H or I-grade wheels would normally be used.

MEASUREMENT OF STRESSES

Stresses were determined by the deflection method. Curvatures parallel and perpendicular to the grinding lines, the directions of the principal stresses (3), were measured as successive uniform layers were etched from the test surface. The principal stresses were calculated by means of the relationship

$$\sigma_1'(w) = \frac{E}{6(1-\nu^2)} \left\{ w^2 \left[\frac{dC_1(w)}{dw} + \nu \frac{dC_2(w)}{dw} \right] + 4w[C_1(w) + \nu C_2(w)] - w_0[C_1(w_0) + \nu C_2(w_0)] - 2 \int_{w_0}^w [C_1(z) + \nu C_2(z)] dz \right\} \dots [1]$$

and a similar equation, with subscripts 1 and 2 interchanged, for $\sigma_2'(w)$.⁵ In Equation [1] w_0 is the initial thickness of the specimen before sectioning and w is the instantaneous thickness as successive layers are removed; C_1 and C_2 are the curvatures in planes containing the normal to the surface and the principal stresses σ_1' and σ_2' , respectively; E is Young's modulus, ν is Poisson's ratio, and z is a co-ordinate parallel to the thickness of the specimen.

Deflection measurements were made with a comparator which utilizes the differential-transformer principle. All measurements were made on 2.0-in. \times 2.0-in. \times 0.120-in. specimens with the exception of those for the bottom surface of bar No. 205, which were made on a 1.25-in. \times 1.25-in. \times 0.120-in. specimen. The former were cut from the bars at position A and the latter at position B in Fig. 1. This was accomplished with an abrasive cutoff wheel mounted on a horizontal spindle surface grinder,

⁵ σ_1' and σ_2' are the principal stresses in a specimen which is constrained so as not to bend. This situation is closely approximated when grinding a workpiece of such rigidity that no appreciable distortion can take place. A summary of the derivation of Equation [1] is given in the Appendix.

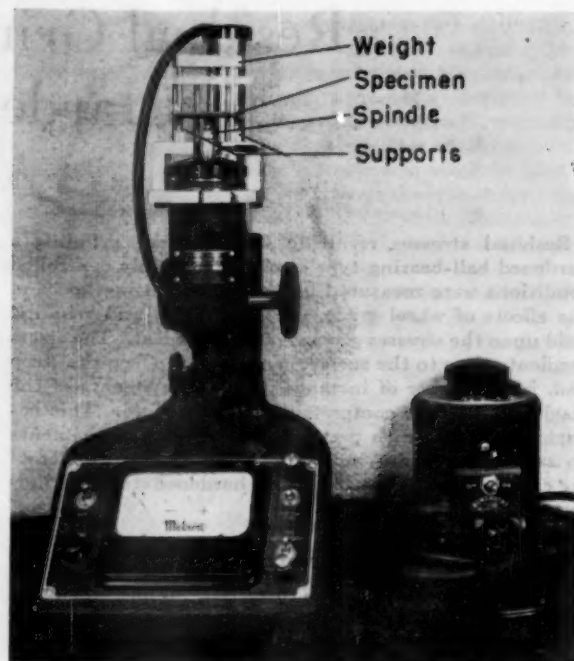


FIG. 2 COMPARATOR WITH 2-IN. SQUARE SPECIMEN IN POSITION FOR MEASUREMENT

using 0.0001-in. downfeeds per traverse and a good flow of grinding fluid to guard against modification of the grinding stresses by heating.

The comparator and a jig to position the specimen accurately for each measurement are shown in Fig. 2. Two supports are so positioned that the spindle of the comparator contacts the reference surface of the specimen, i.e., the surface opposite that being sectioned, midway between the two. A third support, behind the spindle in the photograph, maintains the specimen in a horizontal plane. The spherical tips on the spindle and the two critical supports are of polished boron carbide and have radii of about 0.2 in. A weight with three legs, which rest on the upper surface at points vertically above the supports, prevents the specimen from tipping. The position of the specimen is fixed by three vertical posts which contact two of its edges.

Curvatures along lines through the mid-point of the specimen parallel and perpendicular to the direction of abrasive travel were determined from the displacement of the mid-point of the reference surface with respect to the straight line joining its points of contact with the two spherical boron carbide supports. The supports were spaced 1.8 in. apart for the 2.0-in. square specimens and 1.0 in. apart for the 1.25-in. specimen. Each value of curvature used in plotting working curves for calculating stresses was the average of five separate comparisons of the test specimen with a standard of known flatness.

When the comparator is adjusted to its most sensitive range, each small-scale division corresponds to a 10-microinch displacement of the spindle. In determining the curvatures, readings were estimated to tenths of divisions, a practice justified by the smoothness of the resulting data. The internal consistency of the data is illustrated in Fig. 3 which shows the arc heights obtained as successive layers were etched from the top surface of the specimen cut from bar No. 205.

Both sides of the specimens were sectioned; consequently, a ground surface served as the reference surface when the first side was sectioned and an etched surface when the second side was

sectioned. Subsequent experience has shown that more consistent results are obtained if the reference surface can be lapped or polished. The data in Fig. 3 were obtained with a ground reference surface.

The test surfaces were sectioned by etching in dilute nitric-acid solutions. A 10 per cent solution of concentrated acid in ethyl alcohol was used to remove the first 0.001 in. and a 10 per cent solution in water to remove the remainder of the stressed layer.

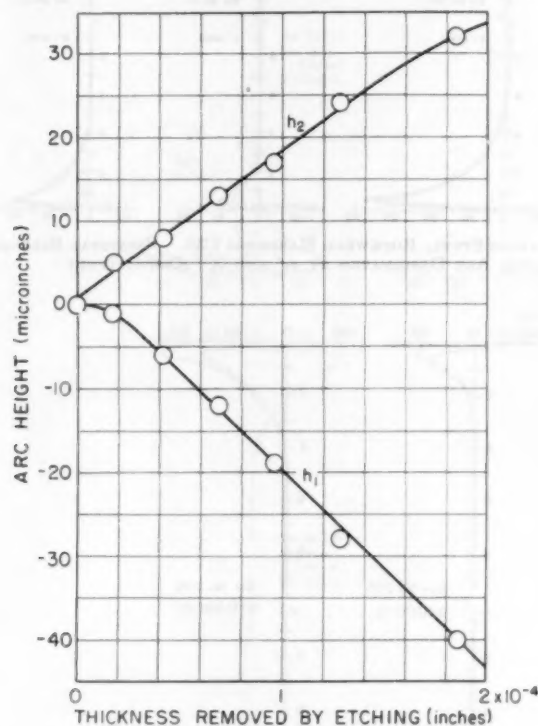


FIG. 3 VARIATION OF ARC HEIGHTS AS SUCCESSIVE LAYERS WERE ETCHED FROM TOP SURFACE OF 2-IN. SQUARE SPECIMEN CUT FROM BAR NO. 205; h_1 AND h_2 REFER TO ARCS PARALLEL AND PERPENDICULAR TO DIRECTION OF ABRASIVE TRAVEL, RESPECTIVELY

Edges of the specimen were protected with a thermosetting resin which remained in place throughout the test. The reference surface was protected with electrical tape which was removed after each etch. The apparatus used for etching is shown in Fig. 4. The specimen was held horizontally by the horseshoe magnet, test surface down, and lowered into the etchant by means of the operating handle until the test surface was immersed. The bath was agitated by the magnetic stirrer which swept away bubbles and improved the uniformity of the etch. The thickness of the layer removed was calculated from the weight loss of the specimen, its area, and density.

RESULTS

In the figures which follow, the principal stresses parallel and perpendicular to the direction of abrasive travel are designated by σ_1' and σ_2' , respectively. Positive values of stress represent tension and negative values compression. All the stress distributions begin 0.00005 in. below the surface because of uncertainties in the behavior of some of the specimens at shallower depths.

Wheel Grade. The stresses generated by grinding with wheels of letter grade H through M at a unit downfeed of 0.001 in., using an oil-water emulsion, are shown in Fig. 5. The maximum tensile stresses in all the surfaces occurred at depths slightly less than

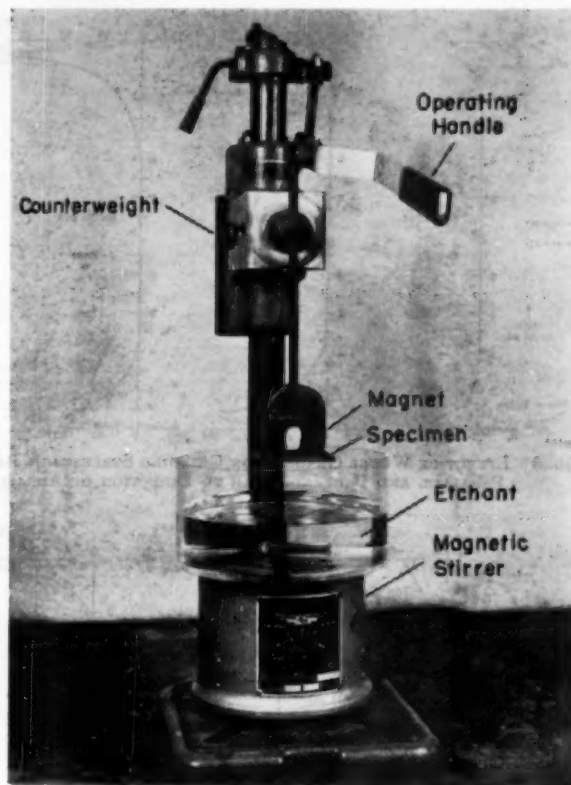


FIG. 4 APPARATUS EMPLOYED IN SECTIONING GROUND SURFACES BY ETCHING

0.001 in. The variation of these peak values with wheel grade was not clean-cut; the difference between the top and bottom surfaces of the same specimen was, in some cases, larger than that between surfaces ground with different wheels. Although the effect of wheel grade is somewhat obscured by the scatter of the results, it is perhaps significant that the highest tensile stresses were found in the bars ground with wheels in the middle of the hardness range tested and that, taking both the top and bottom surfaces into consideration, the smallest and shallowest stresses were found in bar No. 205, ground with the softest wheel tested.

The transverse stress σ_2' was compressive close to the surface in all of the bars represented in Fig. 5. In the bottom surface of bar No. 211, it reached a minimum at a depth of 0.0001 in. and began to increase at shallower depths. The parallel stress σ_1' was tensile, but decreasing, in all except the lower surfaces of Nos. 210, 211, and 212. (The upturn in No. 210 is not shown because it occurred inside the 0.00005-in. minimum depth arbitrarily chosen for plotting the stresses.) The fact that the minima occurred only in the bars ground with the three hardest wheels may be significant. Why they were found only in the bottom surfaces is not clear.

Downfeed. The stresses resulting from grinding with an H-grade wheel at unit downfeeds ranging from 0.0001 to 0.002 in., using an oil-water emulsion, are shown in Fig. 6. In every case the maximum tensile stresses were found a few ten-thousandths of an inch below the surface. Closer to the surface, σ_1' had lower tensile values and σ_2' was compressive for all except the 0.002-in. unit downfeed. Although the lowest maximum tensile stresses were found in the top surface of bar No. 203 and the highest in bar No. 206, the difference was by no means commensurate with

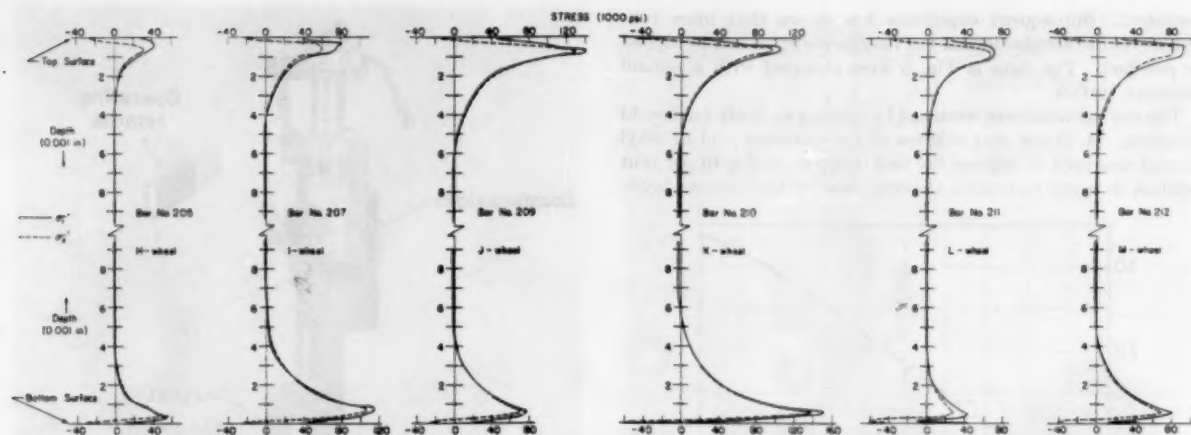


FIG. 5 EFFECT OF WHEEL GRADE UPON GRINDING STRESSES IN BALL-BEARING STEEL, ROCKWELL HARDNESS C59. PRINCIPAL STRESSES PARALLEL AND PERPENDICULAR TO DIRECTION OF ABRASIVE TRAVEL ARE DESIGNATED BY σ_1' AND σ_2' , RESPECTIVELY

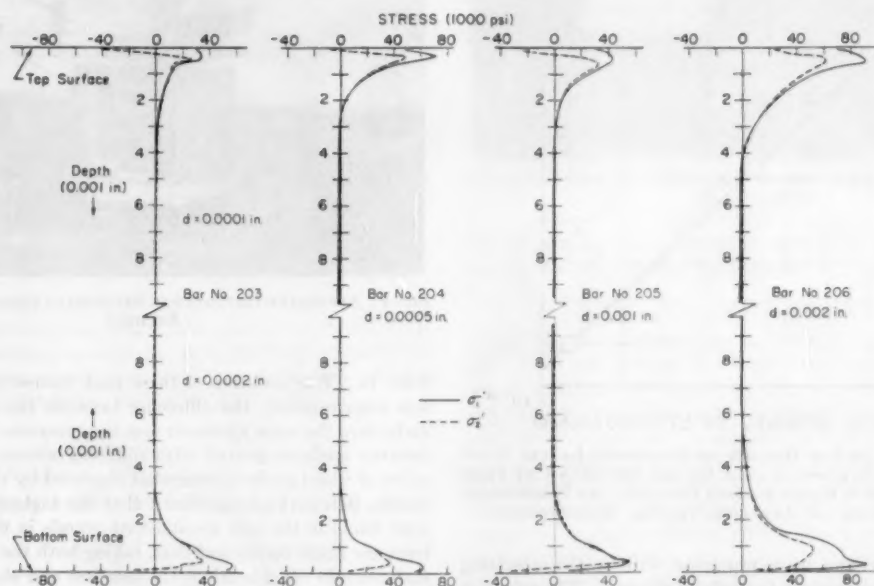


FIG. 6 EFFECT OF UNIT DOWNFEED, d , UPON GRINDING STRESSES IN BALL-BEARING STEEL, ROCKWELL HARDNESS C59. PRINCIPAL STRESSES PARALLEL AND PERPENDICULAR TO DIRECTION OF ABRASIVE TRAVEL ARE DESIGNATED BY σ_1' AND σ_2' , RESPECTIVELY

the twentyfold increase in unit downfeed. In fact, the difference in either the magnitude or depth of penetration of stresses resulting from unit downfeeds up to 0.001 in. was quite small. A unit downfeed of 0.002 in. resulted in a significant increase in both the maximum tensile stresses and the thickness of the stressed layer.

The difference between stresses induced by 0.001 and 0.002-in. unit downfeeds when grinding with an M-grade wheel, using an oil-water emulsion, were somewhat less pronounced than when grinding with the softer H-grade wheel. This behavior is revealed by Fig. 7 which shows that the thickness of the stressed layer was the same in Nos. 212 and 213, ground at downfeeds of 0.001 and 0.002 in., respectively. Although the peak tensile stresses in the top surface of No. 212 were about the same as in both surfaces of No. 213, the average stresses in the latter were somewhat greater. Close to the surface, the stresses in No. 213 were more compressive than in No. 212 and in the bottom surface of both specimens, σ_1' passed through a minimum and

began to increase (algebraically) as it approached the surface.

The effect of downfeed was much more pronounced when using a straight grinding oil with the M-grade wheel, as Fig. 7 also shows. The peak tensile stresses and the thickness of the stressed layer were much smaller in bar No. 215 than in No. 216, ground at unit downfeeds of 0.001 and 0.002 in., respectively. Both specimens exhibited substantial compressive stresses near the surface, about which more will be said later.

Grinding Fluid. The influence of grinding fluid upon residual stresses, touched upon in the preceding section, can be pursued further by again referring to Fig. 7. When using an M-grade wheel at a unit downfeed of 0.001 in., the maximum tensile stresses and the thickness of the stressed layer were much smaller for the straight grinding oil (bar No. 215) than for the emulsion (bar No. 212). At a unit downfeed of 0.002 in., on the other hand, these quantities were higher for the straight oil (bar No. 216) than for the emulsion (bar No. 213).

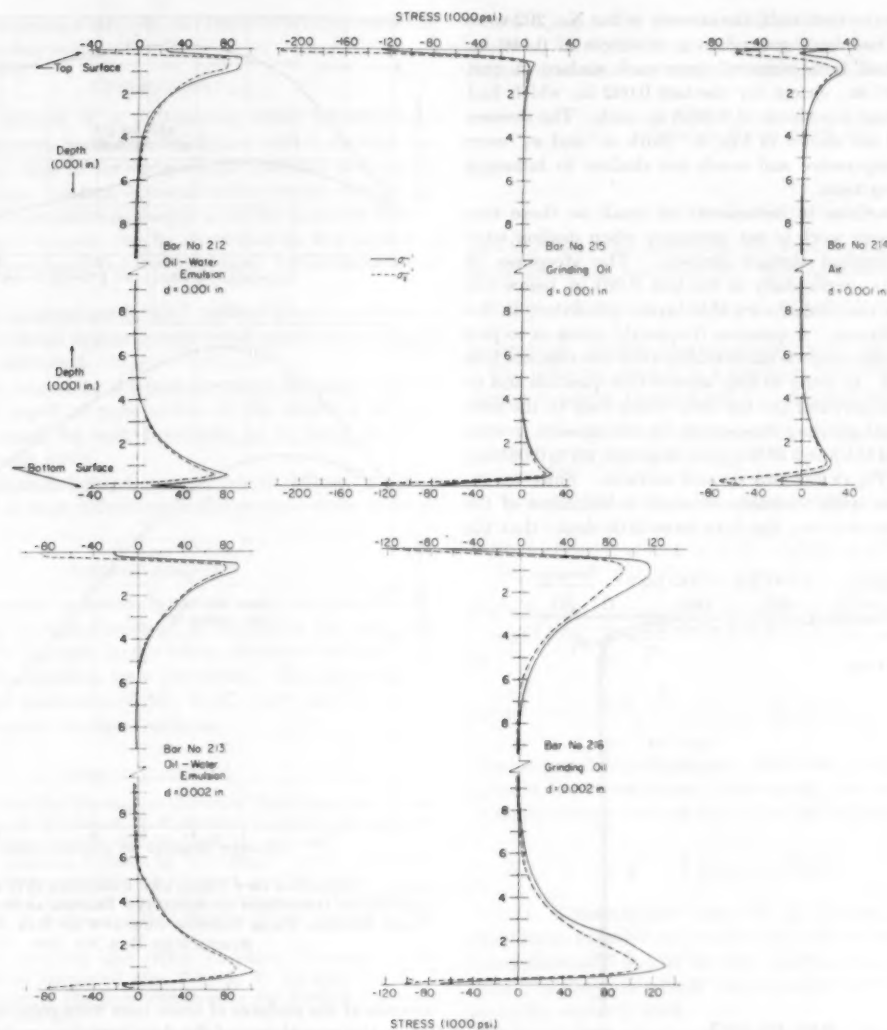


Fig. 7 Effect of Unit Downfeed, d , and Grinding Fluid Upon Residual Stresses in Ball-Bearing Steel, Rockwell Hardness C59, Ground With an M-Grade Wheel. Principal Stresses Parallel and Perpendicular to Direction of Abrasive Travel Are Designated by σ_1' and σ_2' , Respectively

Fig. 7 also provides a comparison of wet grinding with dry grinding when using an M-grade wheel at a unit downfeed of 0.001 in. The stresses induced in bar No. 214, ground dry under conditions otherwise identical with those employed in grinding Nos. 212 and 215, were intermediate in magnitude and depth of penetration between those generated when using the straight grinding oil and the oil-water emulsion. Both σ_1' and σ_2' had minimum values about 0.0002 in. below each surface of bar No. 214 and increased (algebraically) at shallower depths.

DISCUSSION OF RESULTS

The generally good agreement between the stress distributions in the top and bottom surfaces of the bars is indicative of the degree of reproducibility to be expected from separate grinding tests conducted under presumably identical conditions. Discrepancies which appear in Figs. 5, 6, and 7 are believed to be due to statistical variations in grinding conditions beyond the control of the operator. For example, the top surface of bar No. 209 (Fig. 5) exhibited pronounced chatter marks, none of which was found on

the bottom, even though every effort was made to make the grinding conditions the same for both surfaces. This particular bar also affords an interesting sidelight on how wheel chatter may affect residual grinding stresses. The intuitive notion that a chattering wheel peens the surface and tends to introduce compressive stresses is not substantiated by the experimental results. The stresses were more tensile in the top surface, which exhibited chatter marks.

Comparison with previous work (3) reveals that stresses in hardened and tempered ball-bearing steel did not penetrate as deeply as in annealed tool steel of approximately the same carbon content, ground under comparable conditions. Perhaps more important, the stresses in the hardened steel were less tensile at the surface.

The possibility of generating compressive surface stresses by grinding, without inducing tensile stresses of significant magnitude deeper in the metal, was first encountered during the preparation of the specimens. In order to be sure that the stresses resulting from grinding the bars to dimension would not affect the

values obtained in the test itself, the stresses in bar No. 202 were analyzed after it had been ground to a thickness of 0.140 in. About 0.010 in. had been removed from each surface in unit downfeeds of 0.001 in., except for the last 0.002 in. which had been removed in four downfeeds of 0.0005 in. each. The stresses which were found are shown in Fig. 8. Both σ_1' and σ_2' were predominantly compressive and much too shallow to influence subsequent grinding tests.

Sectioning of surfaces in increments as small as those employed in the present work is not necessary when dealing with many types of residual surface stresses. The steepness of the stress gradients, particularly in the first 0.001 in. below the surface, makes the removal of very thin layers mandatory in the case of grinding stresses. A question frequently arises as to just how reliable are measurements made within a few ten-thousandths inch of the surface. In order to help answer this question and to emphasize how unequivocal are the data which lead to the conclusion that residual grinding stresses can be compressive, graphs of curvature versus thickness of the layer removed, up to 0.001 in., are presented in Fig. 9 for two typical surfaces. Since an increase in curvature with thickness removed is indicative of the release of compressive stress, the data leave little doubt that the

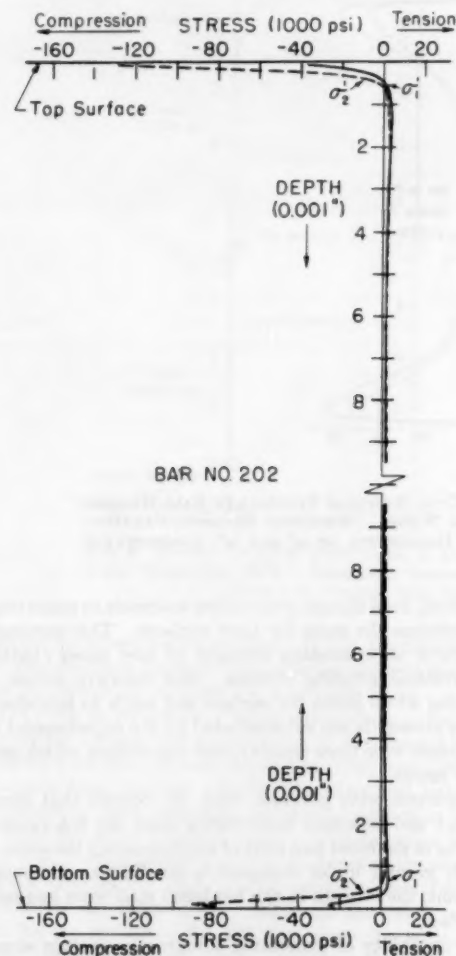


FIG. 8 RESIDUAL STRESSES RESULTING FROM GRINDING BAR NO. 202 TO DIMENSION PRIOR TO TEST; σ_1' AND σ_2' ARE PRINCIPAL STRESSES PARALLEL AND PERPENDICULAR TO DIRECTION OF ABRASIVE TRAVEL, RESPECTIVELY

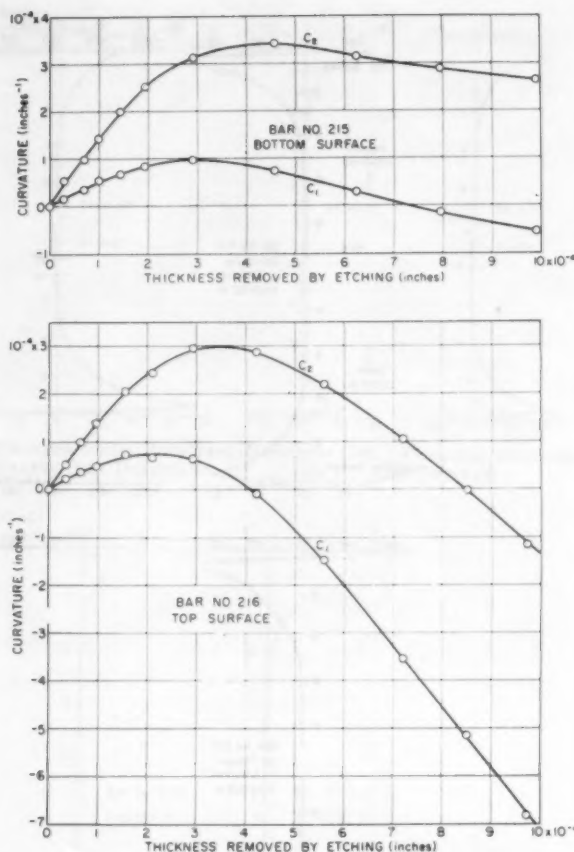


FIG. 9 VARIATION OF CURVATURE PARALLEL (C_1) AND PERPENDICULAR (C_2) TO DIRECTION OF ABRASIVE TRAVEL AS SUCCESSIVE LAYERS WERE ETCHED FROM BOTTOM SURFACE OF BAR NO. 215 AND TOP SURFACE OF BAR NO. 216

stresses at the surfaces of these bars were compressive. Furthermore, the smoothness of the data permits a rather reliable determination of the magnitude of the stresses very close to the surface, certainly up to the 0.00005-in. minimum depth arbitrarily established for calculations in the present work.

A first impulse might be to attribute the high compressive stresses found in some of the bars to rehardening of a thin surface layer, the austenite to martensite transformation causing an expansion of that layer relative to the metal underneath. Taper sections of the surfaces of Nos. 202, 203, 214, 215, and 216 were examined metallographically and no evidence of rehardening was found. It can only be concluded that, if phase transformations of any kind are responsible for the stress behavior, they are not revealed by ordinary metallographic techniques.

The results in Fig. 7 indicate that there are substantial differences between stresses induced by grinding with different grinding fluids; nevertheless, any generalization regarding the effect of a particular type of fluid should be approached with caution. Although certain fluids may be conducive to such behavior, they are not solely responsible for the development of compressive surface stresses by grinding. In an independently conducted test, stress distributions resembling those in bar No. 216 were found in specimens of tool steel, hardened and tempered to Rockwell C55-58, which had been ground dry. This fact suggests that there may be a number of combinations of grinding conditions which

will result in compressive stresses, and research is being continued to discover what they are.

CONCLUSIONS

1 Grinding stresses in a ball-bearing steel, quenched and tempered to Rockwell hardness C59, do not exhibit the high tensile values found close to the surface in an annealed tool steel of comparable carbon content, ground under similar conditions. They also do not penetrate as deeply as in the annealed steel.

2 Compressive stresses can be developed in the surface of quenched and tempered ball-bearing steel by abrasive-wheel grinding.

3 The effect of wheel grade upon residual grinding stresses is small. Very soft wheels appear to reduce the magnitude and depth of penetration somewhat.

4 In general, increasing the unit downfeed increases both the magnitude and depth of penetration of the resulting stresses. The change is small for unit downfeeds up to 0.001 in. when grinding with a soft wheel.

5 Other conditions being the same, use of different grinding fluids can result in large differences in the residual-stress distributions.

ACKNOWLEDGMENT

The author wishes to acknowledge the very considerable contribution of the Norton Company in furnishing the test bars, and of Dr. L. P. Tarasov, under whose direction the bars were prepared and the grinding tests performed. He also wishes to acknowledge the assistance of Mr. R. C. Isler and Mr. J. M. Lesko in carrying out the stress analyses.

BIBLIOGRAPHY

- 1 "Origin of Residual Stresses in Grinding High Chromium Stainless Steels," by L. A. Glickman, T. P. Sanfirova, and V. A. Stepanov, *Journal of Technical Physics*, vol. 19, 1949, pp. 441-447.
- 2 "Residual Grinding Stresses in Mild Steel," by J. Frisch and E. G. Thomsen, *Trans. ASME*, vol. 73, 1941, pp. 337-346.
- 3 "Grinding and Lapping Stresses in Manganese Oil-Hardening Tool Steel," by H. R. Letner and H. J. Snyder, *Trans. ASME*, vol. 75, 1953, pp. 873-882.
- 4 "Effects of Grinding and Other Finishing Processes on the Fatigue Strength of Hardened Steel," by L. P. Tarasov and H. J. Grover, *Proceedings of the American Society for Testing Materials*, vol. 50, 1950, pp. 668-687.
- 5 "Spannungsmessungen an einseitig abgelöschten Knüppeln," by F. Stäblein, *Kruppsche Monatshefte*, vol. 12, 1931, pp. 93-98.
- 6 "Theory of Elasticity," by S. Timoshenko and J. N. Goodier, McGraw-Hill Book Company, Inc., New York, N. Y., second edition, 1951, p. 255.

Appendix

Stäblein (5) showed that, in a bar containing a uniaxial residual stress, the stress at any level w is given by

$$\sigma'(w) = \frac{Ew^2}{6} \frac{dC(w)}{dw} + \frac{2Ew}{3} C(w) - \frac{Ew_0}{6} C(w_0) - \frac{E}{3} \int_w^{w_0} C(z) dz \dots [2]$$

when the bar is held straight by external torques. If the stress is biaxial, each of the principal stresses can be represented by expressions similar to Equation [2]

$$\sigma_1'(w) = \frac{Ew^2}{6} \frac{dC_1(w)}{dw} + \frac{2Ew}{3} C_1(w) - \frac{Ew_0}{6} C_1(w_0) - \frac{E}{3} \int_w^{w_0} C_1(z) dz \dots [3]$$

where $C_1(w)$ is the curvature which $\sigma_1'(w)$, acting alone, would produce. If $C_s(w)$ is the corresponding curvature produced by $\sigma_s'(w)$, the actual curvatures resulting from the combined stresses are

$$C_1(w) = C_{1s}(w) - \nu C_s(w)$$

and

$$C_s(w) = C_{s1}(w) - \nu C_1(w)$$

provided that the deflections are small in comparison with the thickness of the specimen (6). Solving simultaneously

$$C_{1s}(w) = [C_1(w) + \nu C_s(w)] / (1 - \nu^2) \dots [4]$$

is obtained. Substituting into Equation [3] leads to Equation [1].

The principal stresses in a specimen free to bend to its equilibrium curvatures, unrestrained by external forces, are given by

$$\sigma_1(w) = \frac{E}{6(1 - \nu^2)} \left\{ w^2 \left[\frac{dC_1(w)}{dw} + \nu \frac{dC_s(w)}{dw} \right] + 4w[C_1(w) + \nu C_s(w)] + 2(w_0 - 3w)[C_1(w_0) + \nu C_s(w_0)] - 2 \int_w^{w_0} [C_1(z) + \nu C_s(z)] dz \right\} \dots [5]$$

and a similar expression for $\sigma_s(w)$, with subscripts 1 and 2 interchanged (3). The difference between $\sigma_1'(w)$ and $\sigma_1(w)$ is seen to be

$$\sigma_1'(w) - \sigma_1(w) = \frac{E}{1 - \nu^2} \left(w - \frac{w_0}{2} \right) [C_1(w_0) + \nu C_s(w_0)] \dots [6]$$

The principal stresses in the restrained condition thus exceed those in the unrestrained condition by just the stresses induced in straightening the specimen from its initially curved state.

Discussion

A. L. CHRISTENSON⁴ AND W. E. LITTMANN.⁵ This paper contributes valuable information on a subject which has provoked much discussion but all too few experimental results. The high quality of the data reflects the precision which has become familiar in the author's work.

It would be interesting to know whether the abrupt maxima or minima in the stress-distribution curves for cross-feed grinding will also be present in plunge-ground samples. Does the author have any information on stress distribution in plunge-ground-hardened steel? Was cross-feed used in the preparation of the sample of Fig. 8?

X-ray measurements⁷ of surface stress in the direction of grinding have been made on samples carburized and hardened or through-hardened to 60 Rockwell C and ground according to good commercial practice.⁸ The results are similar to those shown in Fig. 8. An interesting observation is that a given grinding condition appears to establish a characteristic surface stress regardless of the sign or magnitude of the prior stress due to heat-treatment. For example, for a prior surface stress of 130,000 psi in compression, grinding produced about 50,000 psi in compression, and for a prior stress of 50,000 psi in tension, grinding produced 30,000 psi in compression.

An over-all view of the stresses produced by grinding leads one to speculate as to their origin. Temperature measurements

⁴ The Timken Roller Bearing Company, Canton, Ohio.

⁷ "X-Ray Measurement of Residual Stress in Hardened High Carbon Steel," by E. S. Rowland and A. L. Christenson, *Trans. ASM*, vol. 45, 1953, p. 638.

⁸ Unpublished research of A. L. Christenson.

within ground samples⁹ show that grinding can produce sufficient nonuniform thermal expansion in the work to cause compressive yielding of the material just beneath the ground surface, thus generating a residual tensile stress. If heating is insufficient to cause yielding, the cold work of the surface in the chip formation may produce the compressive stresses which have been observed by the author and the writers. The fact that similar grinding conditions impose less penetration of tensile stresses in hardened steel than in annealed steel is consistent with the foregoing concept, because a higher peak temperature would be required to cause compressive yielding in hardened steel.

J. FRISCH.¹⁰ The author has presented an interesting study of residual grinding stresses which are found to be of a compressive nature under certain conditions. However, residual stresses, induced in a stress-free material by grinding only, have in the past been shown to be tensile stresses on and immediately below the ground surface.¹¹ The residual stresses which were determined by the author in hardened and ground specimens sectioned from fatigue coupons must be two or more superimposed residual stresses induced by a combination of processes including grinding.

This writer's recent experience in finding considerable compressive residual stresses in high-strength alloy steels heat-treated to various ultimate strength up to 280,000 psi leads one to believe that residual tensile stresses caused by grinding such steels and superimposed on the existing compressive stresses may not necessarily be large enough to nullify them. The finally measured stress values would still be negative, i.e., compressive, but not necessarily pure grinding stresses.

Specimens A and B in Fig. 1 of the paper were removed by means of a cutoff wheel from the hardened and ground fatigue test bars for etching and curvature measurements. Since the residual stresses in manganese oil-hardening tool-steel specimens, which were cut to final shape before hardening and grinding, are reported¹² to be tensile stresses, two questions arise with regard to the different results:

1 Does the use of the abrasive cutoff wheel cause residual stresses in the specimen, similar to those produced by a hacksaw or jeweler's saw,¹³ which may be large enough to influence the original stress patterns considering the 1 to 1 ratio of length to width of the specimens?

2 Does the removal of specimens from fatigue test bars cause them partially to relax before an etchant is applied, and have curvature changes been recorded during and after the use of the cutoff wheel?

Although it is most likely that proper heat-treatment and a unique combination of variables in a subsequent grinding process may produce final residual compressive stresses, the fact that abrasive-wheel grinding can develop compressive stresses would be verified only if either initially stress-free specimens are used, or the exact stress due to a process such as heat-treatment is known.

J. A. MUELLER.¹⁴ It is gratifying to hear that grinding can produce stresses that not only fail to injure the surface but also leave it in a condition that may be desirable from the end use of the ground surface.

⁹ "The Influence of the Grinding Process on the Structure of Hardened Steel," by John Wulff and W. E. Littmann, Trans. ASM, vol. 47, 1955, preprint no. 11.

¹⁰ Assistant Professor of Engineering Design, University of California, Berkeley, Calif. Assoc. Mem. ASME.

¹¹ Reference (2) of the paper.

¹² Reference (3) of the paper.

¹³ "Handbook of Experimental Stress Analysis," by M. Hetenyi, John Wiley & Sons, Inc., New York, N. Y., 1950, p. 466.

¹⁴ The Carborundum Company, Niagara Falls, N. Y.

The pattern of stresses, as found by the author, has a remarkable parallel in the efficiency or grinding ratio of the grinding wheel. The writer would like to bring out this parallel in an attempt to correlate stress with grinding ratio and ultimately, if possible, to set up a means of choosing grinding wheels to produce favorable stresses or eliminate unfavorable stresses without the complexity of lengthy computations.

The author indicates that the stresses produced by a change in wheel grade ranged from a shallow stress in the soft end of the grade range to a maximum in the middle of the hardness range and then to a lesser stress in the hard end. The grinding ratio, we have found, parallels this stress pattern remarkably. Fig. 10 of this discussion shows the grinding ratio of a range of surface-grinding wheels. The soft wheel produces a low grinding ratio, and as the grade of the wheel becomes harder the grinding ratio



FIG. 10 GRINDING RATIO COVERING RANGE OF SURFACE-GRINDING WHEELS

becomes greater until it reaches a maximum and then falls off. In the soft end of the grade range, continuous breakdown occurs and this continues until the wheel becomes too hard for the operation and then discontinuous breakdown sets in. The wheel loads and unloads and breaks out rather than wears away. The grinding ratio strikingly follows the stress pattern as the wheel grade is changed.

In the stress pattern at unit downfeeds ranging from 0.0001 in. to 0.002 in., the author found that there was little difference in the stresses up to a downfeed of 0.001 in. There was a significant change at a downfeed of 0.002 in. Fig. 11, herewith, shows the grinding ratio at various downfeeds. At the smallest downfeed the grinding ratio is small and at heavier infeeds the grinding ratio becomes greater until a maximum is reached and then proceeds to fall off. Again the stress pattern shows a remarkable similarity to the grinding ratio.

The third parallel between stress and grinding ratio occurs when grinding with straight oil, an oil-water emulsion, and air. Grinding oil produced higher stresses than the oil-water emulsion.

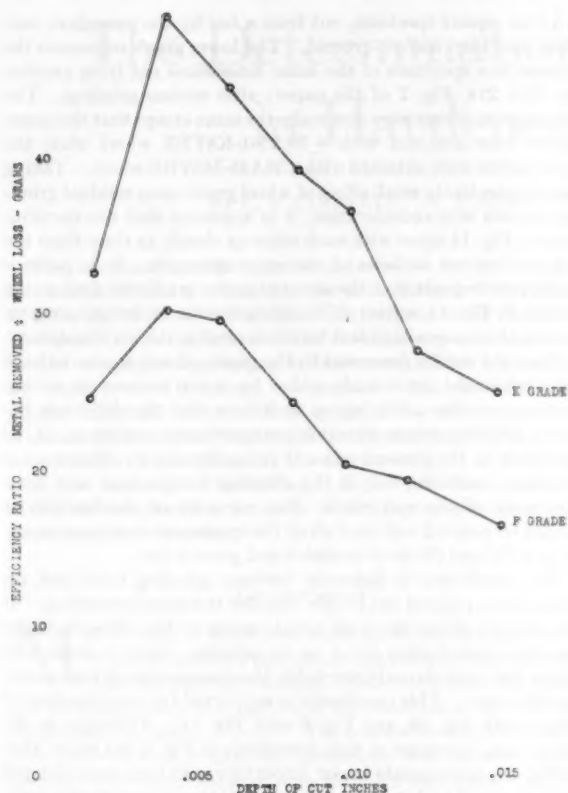


FIG. 11 GRINDING RATIO AT VARIOUS DOWNFEEDS

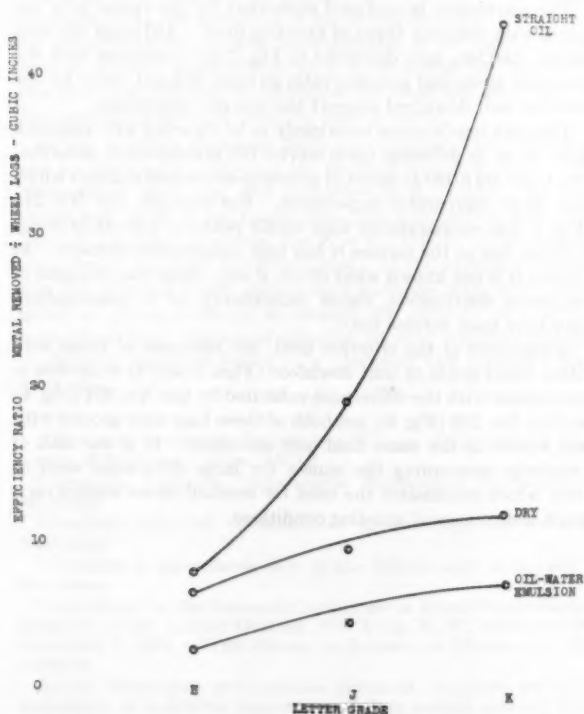


FIG. 12 GRINDING RATIO FOR VARIOUS LUBRICANTS AND DRY

Fig. 12 shows the grinding ratio for straight oil, and oil-water emulsion, and dry. Grinding oil produced the highest grinding ratio, air was the next highest, and the emulsion was the least efficient.

If the author were to have ground in air at the 0.002-in. downfeed the stress pattern might have fallen between the straight oil and the water-emulsion curves.

In summary there is a striking similarity between the stress pattern and the grinding ratio, and it leads to the conclusion that in any grinding operation a compromise must be effected between any or all of the multiplicity of operations that a grinding wheel is called upon to perform. We must compromise among production rate, wheel life, dimensional tolerance, stress generation, form-holding ability, surface finish, stock removal, and many other operations.

AUTHOR'S CLOSURE

The author wishes to thank the discussers for their interesting and helpful contributions to the paper.

The question of cross-feed grinding versus plunge grinding raised by Messrs. Christenson and Littmann is extremely pertinent. The author has no data on plunge-ground surfaces. Cross-feed grinding was normally employed in grinding the test bars to an intermediate thickness of 0.140 in.; however, the exact cross-feed procedure used for the final 0.0005-in. cut on each surface of bar No. 202 (Fig. 8 of the paper) is not known. The fact that such markedly different stress distributions were obtained with a 3-in.-wide wheel suggests that wheel width and cross-feed are important factors for future study.

It is encouraging to learn that another laboratory, using grinding procedures probably somewhat different from those described in the paper and determining stresses by an independent method, finds stress distributions similar to those in Fig. 8. The observation of Messrs. Christenson and Littmann that a given grinding operation establishes a characteristic surface stress, relatively independent of any initial stress resulting from heat-treatment, is substantiated by experience in our own laboratory. In fact, there is a good deal of evidence that each pass of the wheel wipes out existing residual stresses to the depth to which it plastically deforms the surface and introduces its own characteristic stress. If the stress initially present extends deeper than the plastic deformation by grinding, the stress distribution resulting from grinding joins that of the previously existing stress at a depth corresponding to the lower edge of the layer deformed by grinding.

The picture of the origin of grinding stresses proposed by Messrs. Christenson and Littmann is, in general, one to which the author subscribes. The substantial difference between σ_1' and σ_2' near the surface and their rather close agreement at depths below 0.001 in. is consistent with this picture.

Professor Frisch points out that tensile grinding stresses have previously been found on the surface of annealed steel (2, 3) and suggests that the compressive surface stresses found in the present work on hardened steel are a composite of those caused by grinding plus those from some other source, probably heat-treating. Aside from the fact that there appears to be no inherent reason why the grinding stress distributions in quenched and tempered specimens should be identical with those found in annealed specimens of different chemical composition, the suspicion that the results presented in the paper are appreciably affected by initial thermal stresses is not confirmed by experimental data.

Fig. 13 of this closure shows the residual stress in the surface of a typical bar as heat-treated. Recalling that the intermediate grinding removed 0.010 in. and the test grinding another 0.010 in., the initial thermal stress to be considered is that at a depth of 0.020 in. and below in Fig. 13. In this range it does not exceed 1000 psi in absolute value, which is of no significance in com-

parison with the grinding stresses. A further indication of the insignificance of the initial stress is given by Fig. 8 which shows the initial stresses in both surfaces of a typical bar at the beginning of the test grinding. It is quite apparent that the removal of 0.010 in. from each surface gets well below any stress initially present.

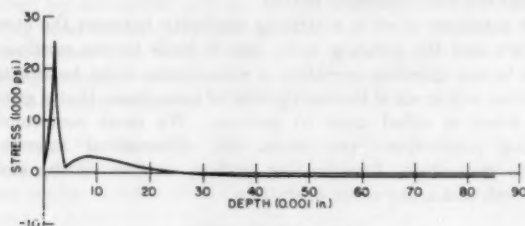


FIG. 13 INITIAL STRESS IN TYPICAL BAR AFTER HEAT-TREATMENT

It does not seem likely that the residual stresses induced in the edges of a specimen by the careful parting procedure described in the paper can influence the determination of stresses in the test surface appreciably. The greatest depth of stress penetration encountered in the surface-grinding experiments, even under the severest conditions tested, was only about 0.008 in. It seems reasonable that stresses set up in the edges of the specimen by the side of the cutoff wheel, operating as described, would be smaller and penetrate less deeply. Irrespective of their magnitude, the fact that they are confined to such a thin layer on the edges practically precludes the possibility of their affecting the determination of stresses in the surface of interest.

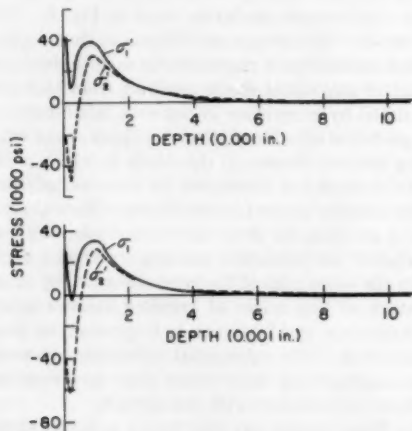


FIG. 14 Above, SPECIMEN PARTED FROM BAR AND THEN SURFACE-GROUND; Below, BAR SURFACE-GROUND AND THEN SPECIMEN-PARTED

This reasoning is substantiated by the graphs in Fig. 14 which also help to answer the question of relief of grinding stresses by the parting operation. The upper graph represents the stresses

in a 2-in. square specimen, cut from a bar by the procedure outlined, and then surface-ground. The lower graph represents the stresses in a specimen of the same dimensions cut from another bar (No. 214, Fig. 7 of the paper) after surface grinding. The grinding conditions were nominally the same except that the upper curves were obtained with a 38A461-K8VBE wheel while the lower curves were obtained with a 38A46-M8VBE wheel. Taking the comparatively small effect of wheel grade upon residual grinding stresses into consideration, it is apparent that the distributions in Fig. 14 agree with each other as closely as those from the top and bottom surfaces of the same specimen. It is perhaps worth pointing out that the steep negative gradients close to the surface in Fig. 14 appear to be characteristic of dry grinding on steel of the composition and hardness used in the experiments.

Since the results presented in the paper do not appear to have been influenced appreciably either by initial stresses or by the parting procedure, it is logical to believe that the difference between grinding-stress distributions previously published (2, 3) and those in the present paper is primarily due to differences in grinding conditions and in the chemical composition and heat-treatment of the specimens. For purposes of clarification, it should be pointed out that all of the specimens discussed in references (2) and (3) were annealed and ground dry.

The parallelism in behavior between grinding ratio and residual stress pointed out by Mr. Mueller is indeed interesting. If one chooses either the peak tensile stress or the net area under the stress-distribution curve as his criterion, there is some indication that high stress is inevitably the consequence of high grinding efficiency. This conclusion is supported by a comparison of Fig. 5 with Fig. 10, and Fig. 6 with Fig. 11. Although, in the latter case, the range of unit downfeeds in Fig. 6 lies below that in Fig. 11, experiments in our laboratory with both annealed and fully hardened tool steel have shown that the area under the stress curve increases with unit downfeed and reaches a maximum in the neighborhood of 0.004 to 0.006 in.

The correlation is confused somewhat by the stress data obtained with different types of grinding fluid. Although the data for the 0.002-in. unit downfeed in Fig. 7 are consistent with the view that stress and grinding ratio go hand in hand, those for the 0.001-in. unit downfeed support the opposite conclusion.

The peak tensile stress is certainly to be regarded with suspicion in so far as its influence upon service life is concerned; nevertheless, there are other features of grinding-stress distributions which may be of comparable importance. For example, bar No. 216 (Fig. 7) has comparatively high tensile peaks at a depth of about 0.001 in. but at the surface it has high compressive stresses. At present it is not known what effect, if any, these two features of the stress distribution, either individually or in combination, may have upon service life.

Irrespective of the criterion used, the variation of stress with either wheel grade or unit downfeed (Figs. 5 and 6) is modest in comparison with the differences exhibited by bar No. 204 (Fig. 6) and bar No. 202 (Fig. 8), yet both of these bars were ground with soft wheels at the same final unit downfeed. It is our lack of knowledge concerning the causes for large differences such as these which emphasizes the need for residual stress studies on a much wider range of grinding conditions.

The Determination of Residual Stresses in Hardened, Ground Steel

By L. V. COLWELL,¹ M. J. SINNOTT,² AND J. C. TOBIN,³ ANN ARBOR, MICH.

The residual surface stresses induced by grinding a hardened SAE 4340 steel have been investigated by means of x-ray diffraction and by optical interferometric methods. The depth of penetration of the residual stresses increases as the severity of grinding is increased. The higher the hardness level of the steel, the greater is the absolute value of the residual stress. The magnitude of the residual stresses varies from 100,000 psi in tension to 135,000 psi in compression depending on the hardness level of the steel and the severity of grinding. Light grinding produces residual compressive stresses while heavy grinding produces tensile stresses.

INTRODUCTION

MANY studies have been made on the stresses induced in steels by grinding operations and it is generally accepted that the grinding produces tensile stresses in the surface and their magnitude is a function of the condition of the steel and the grinding variables (1 to 5).⁴ The present investigation differs from previous work in that two methods for residual-stress measurement were used; steels of high hardness were used; grinding more closely approximated industrial practice in that multiple-pass grinding was used.

The higher-strength steels, i.e., those heat-treated to the higher hardness levels, are becoming of increasing industrial significance for many applications, and since grinding is the most common method of finishing machined parts at these high hardness levels this work was initiated to determine the sense and magnitude of the residual stresses produced by such grinding.

In any grinding operation there are numerous variables which can be expected to influence the induced surface stresses. The composition, structure, and properties of the metal; the feed, speed, and depth of cut in grinding; the composition, grain, and bond of the grinding wheels; the type and quality of the coolant; are a few of the more important ones. In the present investigation the grinding wheel and coolant variables were not investigated or changed except to change the severity of grinding. Several arbitrary degrees of grinding severity were investigated at different hardness levels on a quenched and tempered SAE 4340 steel.

¹ Professor, Production Engineering, College of Engineering, University of Michigan. Mem. ASME.

² Associate Professor, Metallurgical Engineering, University of Michigan.

³ Graduate Student, Metallurgical Engineering, University of Michigan.

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 2, 1954. Paper No. 54-A-52.

RESIDUAL-STRESS-MEASUREMENT METHODS

The use of x-ray diffraction for stress measurements is not new but its use on hardened steels has not been too successful because of several complicating factors. The use of x-rays is based on the fact that residual stresses cause distortion of the metal lattice and this distortion is reflected in changes in interplanar spacings which can be detected by x-ray diffraction. The differentiated Bragg equation gives

$$d(d)/d\theta = -\frac{\lambda \cot \theta}{2 \sin \theta}$$

where

d = interplanar spacing

θ = diffraction angle

λ = wave length

The nature of the geometry of the high-angle spectrometer is such that the d -value in the foregoing equation is for planes which lie parallel to the specimen surface being irradiated with x-rays. The measurements are made perpendicular to the surface while the stresses to be evaluated are parallel to this surface. Elastic theory is used to relate these stresses as follows

$$\epsilon_z = \Delta d/d; \quad \sigma_{x,y,z}/\epsilon_{x,y,z} = E; \quad \nu = -\epsilon_x/\epsilon_x + \epsilon_y = -E\epsilon_x/\sigma_x + \sigma_y; \quad \Delta d/d = -\nu\sigma_x/E = -\cot \theta \Delta\theta; \quad \sigma_x = \sigma_x + \sigma_y; \quad \sigma_z = E\Delta\theta/\nu \tan \theta$$

where

$\epsilon_{x,y,z}$ = strain in x , y , or z -direction

$\sigma_{x,y,z}$ = stress in x , y , or z -direction

E = Young's modulus, 30×10^6 psi

ν = Poisson's ratio, 0.28

From the Bragg equation it is apparent that the maximum sensitivity for the determination of the residual stress is approached in the back-reflection region of x-ray diffraction. The choice of radiation will determine how far back into this region it is possible to go since the material to be investigated has fixed parameters. Other factors such as absorption, fluorescence, scattering, etc., must be taken into account and this then limits the radiation to either chromium or cobalt. Cobalt was used in this work since it gave the greatest possible angle of diffraction with the unit available although chromium radiation would have been better in that a greater intensity of diffraction would be obtained. A Norelco high-angle spectrometer was used in this work.

While the foregoing are the essential theoretical considerations, there are several complicating factors which must be taken into account. Since the material being investigated is a hardened steel, considerable internal stress is already present before grinding, and the effect of this stress on the x-ray pattern is to produce broad diffuse lines in place of the sharp well-defined lines that would be obtained from a stress-free metal. This broadening of the diffraction lines causes merging of the $\alpha_1\alpha_2$ doublet and in order to separate this doublet an extensive correction technique must be applied to the experimentally determined peaks. The separation of the doublet is accomplished by a method suggested by Finch (6). This technique assumes that the composite in-

tensity of the diffracted beam can be represented by the relationship

$$I = I\alpha_1 + I\alpha_2 = \text{sech}^2 hx + \frac{1}{2} \text{sech}^2 h(x-1)$$

where

I = intensity

h = factor associated with line width and a function of stress, faulting, fragmentation, etc.

x = angular units wherein the completely resolved doublet separation is equal to unit (i.e., α_1 peak $x = 0$, α_2 peak $x = 1$)

Since only the peak position of the α_1 line is desired, the integrated line intensity obtained experimentally must be corrected for the apparent shift caused by the α_2 line. This correction can be calculated from the intensity equation by plotting this equation for various values of h . From this plot a first estimate of the apparent peak is determined. The peak position is calculated by differentiating the intensity equation to give

$$\tanh(hx) \text{sech}^2(hx) = -\frac{1}{2} \tanh h(x-1) \text{sech}^2 h(x-1)$$

and solving this graphically. The substitution of the peak value into the previous equation gives the half-height of the line and this then permits the graphical determination of the width.

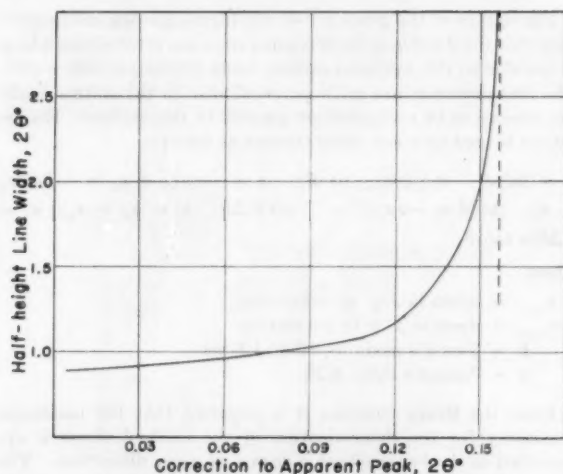


FIG. 1 SHIFT OF APPARENT PEAK INTENSITY DUE TO MERGING OF DOUBLET

Fig. 1 is a plot constructed as indicated in the foregoing. It is used by plotting the actual line contour, measuring the half-height width, using Fig. 1 to obtain the correction, and this is then subtracted from the apparent peak as determined by least-squares analysis. In determining the peak intensity of the trace of the 220 line of the ferrite in the steel, a manual counting technique was used. A constant count procedure was used at increments of 0.02 deg 2θ and the resulting trace of the curve plotted and fitted to an equation using the method of least squares. This equation was then differentiated to find the maximum. If carried out manually, this procedure is extremely laborious but by using a CPC-IBM computer these data can be quickly and accurately reduced.

The interferometric method used to study the residual stresses was identical

to that described by Letner (5). Briefly, the method involves following the change in curvature of an initially flat work specimen as successive layers of the stressed surface are etched away. Calculations, involving the change of curvature plus a correction for elastic relaxation following etching, serve as the basis for determining the continuous stress-distribution function for both principal stresses in the plane of the test plate. Note that this is different from the method used in x-ray analyses which gives only a single value, the stress sum for the plane of the test plate. The stress functions obtained in the interferometric method are continuous despite differential etching since the principal stress component is based on the rate of change of curvature as measured from continuous plots of curvature versus the amount of metal etched from the stressed surface.

The equations for the principal stresses and a sketch of the dimensional properties of a specimen are shown in Fig. 2. In all the grinding operations one principal stress was parallel to the cutting direction. The meaning of the symbols in Fig. 2 is as follows:

- σ_1 and σ_2 = principal stresses in planes parallel to test surface
- $w_0 - z$ = cumulative thickness of metal removed by etching
- w_0 = original thickness of material
- E = Young's modulus
- ν = Poisson's ratio
- C_1 = curvature of specimen in direction of σ_1 after etch
- C_2 = curvature of specimen in direction of σ_2 after etch

EXPERIMENTAL PROCEDURES

The steel used in this work was taken from a billet of SAE 4340 steel 2 in. \times 4 in. \times 16 in. This billet was sectioned to produce pieces $\frac{7}{8}$ in. \times $\frac{7}{8}$ in. \times $\frac{1}{16}$ in. and 2 in. \times 2 in. \times $\frac{1}{16}$ in. The pieces were hardened by heating to 1550 F for 2 hr, then quenching into oil. Tempering was carried out by heating to temperatures of 300, 450, 700, and 1100 F for 1 hr in a recirculating air furnace. These treatments produced hardness values of RC-55, RC-50, RC-46, and RC-32, respectively. Any decarburization that may have occurred was removed by grinding and lapping the test specimens. After lapping, the specimens were etched in a 5 per cent nital solution to remove any residual stresses put into the surface by the prior grinding or lapping operations. Approximately 0.006 in. was removed in this manner. This procedure has been shown to be effective in removing the disturbed surface without introducing any additional stresses (7). All grinding was parallel to billet length.

After this treatment the 2-in. \times 2-in. \times $\frac{1}{16}$ -in. pieces were metallographically polished on one of their flat sides to produce a specular surface. This surface was maintained throughout all subsequent treatments in order to follow the change in residual stress by the optical interferometric method described previously. The smaller pieces were used for the x-ray diffraction analyses.

The test specimens were then carefully ground under the conditions listed in Table 1. Five consecutive passes were made over the surface of each piece. Conditions 1, 2, and 6 at hardness levels of RC-55, RC-50, and RC-46 were examined for residual stress by the x-ray diffraction method while all conditions at all hardness levels were examined by the interferometric method (5, 7, 8).

The ground specimens for x-ray analysis were mounted in a special fixture to position the pieces in exactly the same position

TABLE 1 GRINDING CONDITIONS

Condition no.	Depth of cut, in.	Feed, in.	Cutting speed, fpm	Work speed, fpm	Wheel diam and width, in.	Wheel material
1.....	0.0007	0.050	6000	45	12 \times 1	32A46J8VBE
2.....	0.0007	0.050	4000	30	8 \times 1	32A46J8VBE
3.....	0.0007	0.050	2000	15	4 \times 1	32A46J8VBE
4.....	0.0007	0.050	6000	45	12 \times 1	32A80J8VBE
5.....	0.0015	0.050	6000	45	12 \times 1	32A80J8VBE
6.....	0.0015	0.050	6000	45	12 \times 1	32A80J8VBE

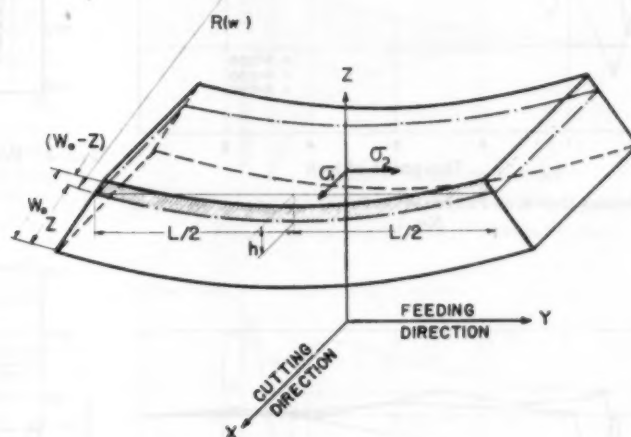
NOTE: Coolant was 5 per cent NaNO₂ by weight except for Condition 6 which was dry.

$$\sigma_1(z, w_0) = \frac{E}{6(1-\nu^2)} \left[z^2 \left[\frac{dC_1(z)}{dz} + \nu \frac{dC_2(z)}{dz} \right] + 4z \left[C_1(z) + \nu C_2(z) \right] + 2(w_0 - 3z) \left[C_1(w_0) + \nu C_2(w_0) \right] - 2 \int_z^{w_0} [C_1(z) + \nu C_2(z)] dz \right]$$

$$\sigma_2(z, w_0) = \frac{E}{6(1-\nu^2)} \left[z^2 \left[\frac{dC_2(z)}{dz} + \nu \frac{dC_1(z)}{dz} \right] + 4z \left[C_2(z) + \nu C_1(z) \right] + 2(w_0 - 3z) \left[C_2(w_0) + \nu C_1(w_0) \right] - 2 \int_z^{w_0} [C_2(z) + \nu C_1(z)] dz \right]$$

$$C = \frac{8h}{L^2}$$

FIG. 2 GEOMETRY AND NOMENCLATURE OF LETNER INTERFEROMETRIC METHOD OF ANALYSIS



in the x-ray beam and a trace of the 220 ferrite line was obtained, the peak position determined and corrected as previously described. The specimen was then coated with paraffin except on the ground surface and a layer etched off this surface with nital for a depth of a few ten thousandths of an inch. The specimen was then cleaned and remounted in the x-ray fixture and the position of the peak redetermined. Successive layers were removed and the peaks determined until no change in the location of the peak was detected. This condition was then taken as the stress-free reference point and using these data it is possible to compute the stress at any point referred to this base using the relationship

$$S_x = s_x + s_y = (\theta - \theta_0)E/\nu \tan \theta$$

where

θ = diffraction angle where stress is to be determined

θ_0 = diffraction angle of stress-free reference point

E = Young's modulus, 30×10^6 psi

ν = Poisson's ratio, 0.28

Figs. 3, 4, and 5 summarize the data obtained in this fashion for

with paraffin except for the ground surface which was incrementally etched in steps of 0.00005 to 0.001 in.; the smaller increments were taken near the work surface. The curvatures of the specimen were measured in the principal stress directions after each etch. The resulting curvature-depth curves provided the basis for the calculation of the principal stresses parallel to the surface using the equations shown in Fig. 2. Typical plots of the principal stresses versus depth below the surface are shown in Figs. 6, 7, and 8.

A more complete summary of the stress condition of the various test specimens is given in Table 2. The entries in this table are in the form of a numerator and a denominator. The numerator is the curvature multiplied by 10^4 and the denominator is the total depth of affected or stressed metal expressed in thousandths of an inch. Increased or positive curvature change upon removal of metal by etching indicates compressive stress while negative curvature change indicates tensile stress. The ratio of the numerator to the denominator is a gross indication of the predominant stress since the principal component of the stress function is linearly proportional to the rate of change or the derivative of the

TABLE 2 SUMMARY OF GROSS CURVATURE CHANGE IN RESIDUAL-STRESS SPECIMENS

Grinding operation no.	Curvature change/affected depth = in. $\times 10^4$ /in. $\times 10^3$							
	RC-55 Cutting, feeding		RC-50 Cutting, feeding		RC-46 Cutting, feeding		RC-32 Cutting, feeding	
1.....	-2.4/7	-1.2/7	-0.7/4	-0.8/5	-0.3/2	+1.3/2	+0.1/2	+1.7/2
2.....	-1.7/6	-2.0/6	-0.3/3	+1.0/3	-0.4/3	+1.5/3	+0.4/3	+1.5/3
3.....	-1.2/6	-1.5/6	+0.5/2	+1.2/2	+0.3/2	+1.2/2	+0.4/3	+1.2/2
4.....	-3.0/4	-1.2/4	-0.3/4	+0.3/4	-0.3/4	+0.5/4	-1.2/4	+0.3/4
5.....	-5.2/9	-2.5/9	-0.5/3	-0.6/3	-0.5/3	+1.0/3	-0.2/3	+1.2/3
6.....	-3.0/6	-1.4/6	-0.7/6	+0.8/6	-3.5/6	-1.3/6	-3.2/6	-0.7/6

the steel at three hardness levels which had been ground under three different conditions.

The test specimens for the interferometric analysis were treated in a similar manner to the x-ray specimens. They were coated

curvature with respect to the thickness of the affected metal. The entries of Table 2 reflect the severity of the grinding operations in the following ways: The minimum average stress is proportional to the ratios given; unit ratio corresponds to a stress of

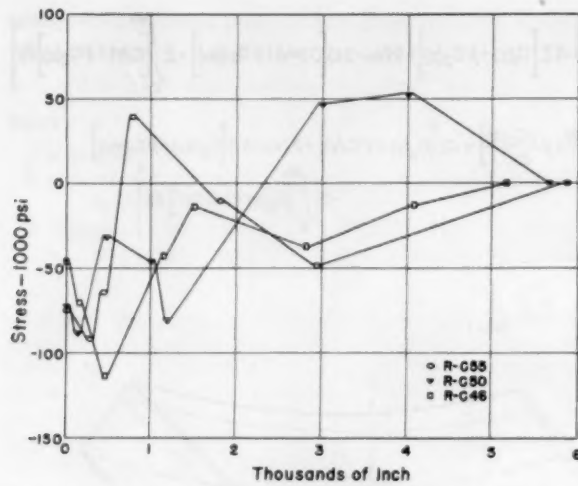


FIG. 3 STRESS DEPTH OF PENETRATION PLOT—GRINDING CONDITION No. 1

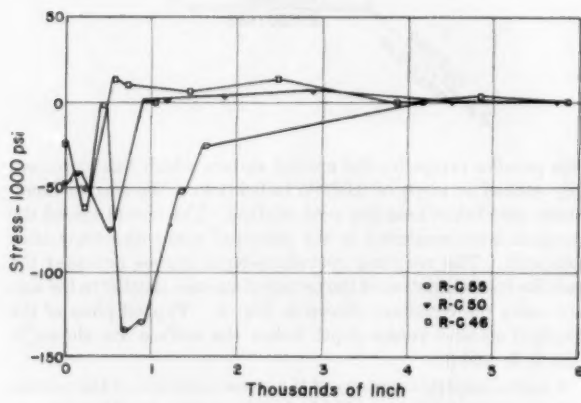


FIG. 4 STRESS DEPTH OF PENETRATION PLOT—GRINDING CONDITION No. 2

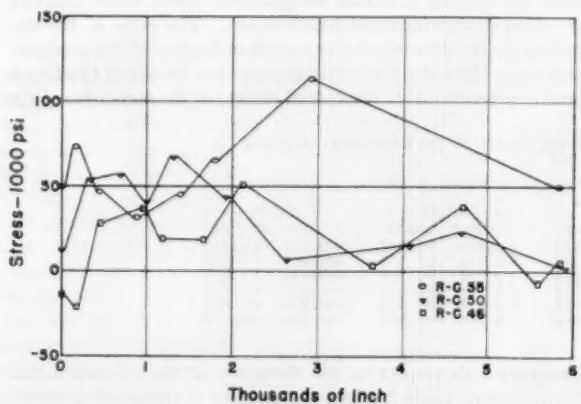


FIG. 5 STRESS DEPTH OF PENETRATION PLOT—GRINDING CONDITION No. 6

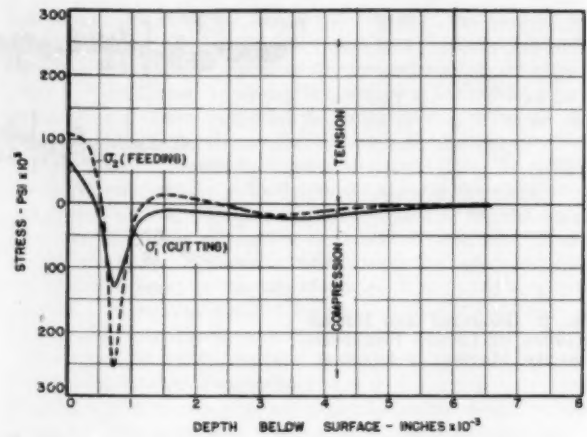


FIG. 6 PLOT OF PRINCIPAL STRESSES VERSUS PENETRATION—GRINDING CONDITION No. 1, RC-55

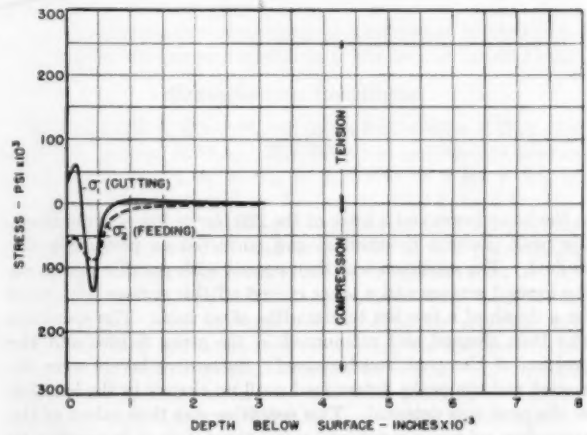


FIG. 7 PLOT OF PRINCIPAL STRESSES VERSUS PENETRATION—GRINDING CONDITION No. 1, RC-46

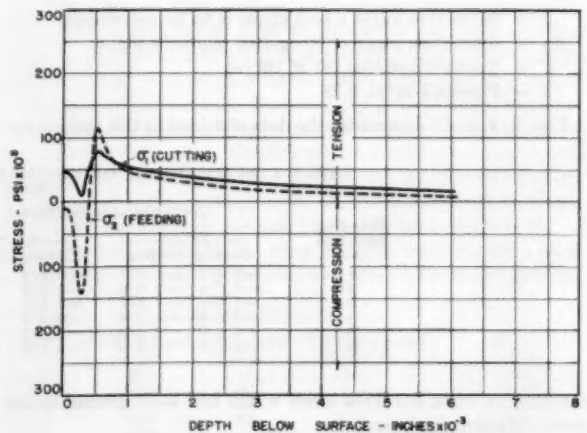


FIG. 8 PLOT OF PRINCIPAL STRESSES VERSUS PENETRATION—GRINDING CONDITION No. 5, RC-55

about 45,000 psi. The denominators indicate the depth of the affected region. The absolute value of the numerator is a measure of the energy absorbed through strain distortion.

Table 3 contains a set of values which are a rough measure of the type and extent of the strain disturbance. These values were obtained by the algebraic addition of the numerators for the cutting and feeding directions shown in Table 2. Table 4 contains a set of values which are a rough measure of the energy absorbed by the strained surfaces. They are the product of the average de-

TABLE 3 SUMMARY OF STRAIN

Operation no.	(Sum of numerators in Table 2)			
	55	Work hardness, RC		
		50	46	32
1.....	-3.6	-1.5	+1.0	+1.8
2.....	-3.7	+0.7	+1.1	+1.9
3.....	-3.7	+1.7	+1.5	+1.6
4.....	-4.2	0	+0.2	-0.9
5.....	-7.7	-1.1	+0.5	+1.0
6.....	-4.4	+0.1	-4.8	-3.9

TABLE 4 COMPARISON OF ENERGY ABSORPTION

Operation no.	(Product of average depth and absolute value of curvature change)			
	55	Work hardness, RC		
		50	46	32
1.....	25.2	6.7	3.2	3.6
2.....	22.2	3.9	5.7	5.7
3.....	22.2	3.4	3.0	4.0
4.....	16.8	2.4	3.2	6.0
5.....	69.3	3.3	4.5	4.2
6.....	26.4	9.0	29.0	23.4

nominators and the sum of the absolute values of the numerators of Table 2. The data of both of these tables show that tensile stresses dominate in all grinding operations on the RC-55 specimens, that tensile stresses dominate in all of the dry grinding specimens, that the dominant stress is slightly on the compressive side for all other conditions, and that there is no appreciable difference between these various surfaces.

SOURCES OF RESIDUAL STRESS

It is well known that the cutting forces in grinding operations are considerably smaller than those peculiar to the "thick-chip" operations such as milling or shaping. There is also a significant difference in the ratio of the cutting component to the normal or thrust component. The ratio is usually about 1:2 for the thick-chip operations and 2:1 for grinding operations. It is also known that the rate of energy release per unit volume of metal removed is from 5 to 20 times as great for grinding as compared to other methods of metal removal. These facts in combination with the stress-distribution curves obtained in this study would lead one to classify the stress sources as being either mechanical or thermal in nature or some combination of these two. If one associates the compressive stresses with mechanical sources, and tensile stresses with the thermal sources then the shape of the stress-distribution curves can be considered as the sum of these two effects.

Fig. 9 illustrates the sense and magnitude of these stress sources, while the plots of Fig. 10 show how these sources might superimpose in the actual grinding process. The mechanical and thermal sources of stress will differ as to magnitude and depth of penetration, depending on the grinding conditions, and the dominance of one or the other will be reflected in the resulting stress distribution. Fig. 10 illustrates this combination of effects. Case I shows an operation in which the thermal source is dominant while Case II is one in which the thermal effect is smaller in comparison to the mechanical effect. The MC refers to mechanical cutting; the TC to thermal cutting; MR to mechanical rubbing; and TR to thermal rubbing.

LIMITATIONS ON METHODS

The x-ray method of determining residual stress as used in this

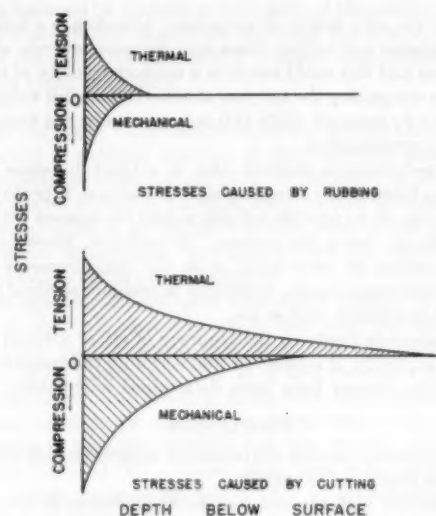


FIG. 9 SCHEMATIC SKETCH OF STRESS SOURCES

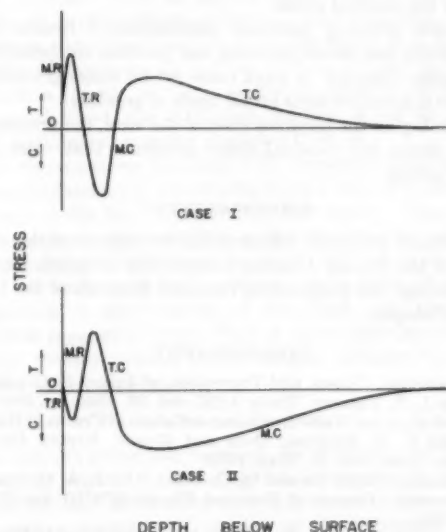


FIG. 10 SCHEMATIC STRESS-DISTRIBUTION CURVES

investigation has several limitations which should be recognized. Cobalt radiation and the 220 line of ferrite were used. This gives a 2θ -value on the order of 124 deg which is not far enough back in the back-reflection region for maximum sensitivity. The use of Cr radiation would have placed this angle at 156 deg and, in addition, would have produced a more intense line for analysis. Since the 220 line of the ferrite was used it should be recognized that changes in the ferrite due to tempering may occur from the heating effect produced in grinding. The quenched and tempered steel is not in an equilibrium state and the grinding heat may have permitted some precipitation of the carbide from solution and this could change the interplanar spacings that were measured. The nonhomogeneity of the grinding operation and the structure of the steel might contribute considerable uncertainty to the stress determination.

While duplicate tests were made and fair agreement obtained, duplicates checked to within ± 4000 psi, there is some scatter from specimen to specimen. Some of the scatter is statistical but

this is not the only source of variation. Grinding is a nonhomogeneous process and certain areas may be more severely abraded than others and this could result in a nonhomogeneity of residual stress. In computing the residual stresses the moduli values were assumed to be isotropic while this is known to be not true except as a first approximation.

The interferometric method also is subject to some of the limitations listed in the foregoing and, in addition, is probably not sensitive enough to provide reliable values for stresses within the first 0.0005 in. below the surface. In addition, there are other possible sources of error such as in the measurements of the optical interference bands, variations in etching, residual stresses from heat-treatment, and so on.

While absolute limits on the accuracy of either method cannot be stated explicitly, it is believed that certainly the magnitude and sense of the stresses have been determined adequately.

CONCLUSIONS

1 All grinding stresses are relatively superficial, not exceeding 0.006 in. in depth on the average.

2 Critically high stresses and stress gradients do not exceed 0.002 to 0.003 in. in depth below the ground surfaces.

3 The greater the severity of grinding, the greater is the magnitude of the residual stress.

4 Severe grinding produces predominantly tensile surface stresses while less severe grinding can produce compressive surface stresses although in most cases severe stress gradients can exist near the surface layer for all kinds of grinding.

5 The higher the hardness level of the steel, the greater is the residual stress and residual-stress gradients that exist in the ground surface.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of this investigation by the Bendix Aviation Corporation of South Bend, Indiana, through the Engineering Research Institute of the University of Michigan.

BIBLIOGRAPHY

- 1 "Detection, Causes, and Prevention of Injury in Ground Surfaces," by L. P. Tarasov, Trans. ASM, vol. 36, 1946, pp. 389-440.
- 2 "Grinding and Heat-Treatment as Causes of Cracks in Hardened Steels," by C. E. Sweetser, *Grits and Grinds*, Norton Company, Worcester, Mass., vol. 21, May, 1930.
- 3 "Residual Stress Caused by Grinding," by L. A. Grikman and V. A. Stepanov, *Journal of Technical Physics (USSR)*, vol. 10, 1946, pp. 791-802.
- 4 "Residual Grinding Stresses in Mild Steel," by J. Frisch and E. G. Thomsen, Trans. ASME, vol. 73, 1951, pp. 337-346.
- 5 "Application of Optical Interference to the Study of Residual Surface Stress," by H. R. Letner, *Proceedings of the Society for Experimental Stress Analysis*, vol. 10, no 2, pp. 23-36.
- 6 "A Correction to the Diameter Measurement of Diffuse X-Ray Diffraction Rings," by L. G. Finch, *Nature*, vol. 163, no. 4141, 1944, p. 32.
- 7 "The Effect of Quenching and Tempering on Residual Stresses in Manganese Oil Hardening Tool Steel," by H. J. Snyder, Trans. ASM, vol. 45, 1953, pp. 605-619.
- 8 "Grinding and Lapping Stresses in Manganese Oil-Hardening Tool Steel," by H. R. Letner and H. J. Snyder, Trans. ASME, vol. 75, 1953, pp. 873-882.

Discussion

J. L. BEATON.¹ The necessity for evaluation of effects of grinding has of recent years assumed ever-larger proportions. Microstructural changes of rehardening and retempering during abusive grinding have received attention in view of their demon-

strated deleterious effect on cutting-tool life. Further than this, applications involving necessity for electroplating subsequent to grinding operations, have resulted in demonstrations of high-order residual stresses in superficial layers of ground surfaces when additive stress-corrosion or hydrogen-embrittling effects of the electroplating operation or preparatory chemical cleaning have resulted in cracking of the steel beneath the plate. While metallurgical transformations of steel surfaces during grinding are probably the worst offenders in promoting high-order stresses that result in such "grinding checks," introduction of stresses by abusive grinding short of resulting in transformation may be of an order sufficient to result in the same effect. Indeed, supplemented by inherently introduced tensile stresses of certain electroplates such as chromium, tensile stresses induced in grinding may easily tip the balance toward premature fatigue failures even though cracking may not have resulted during the plating operation itself.

Because of serious economic factors in loss of material in finished stages of fabrication as well as possible structural inadequacies in service involving fatigue, it was obvious that a tool capable of evaluating grinding practices for their degree of effect on structural behavior needed to be developed. It is believed that the authors have applied techniques to the problem that have demonstrated their applicability by confirmation of theoretically derived data as well as certain effects noted in large-scale production of electroplated high-strength steel cylinders. Certainly the studies have not provided the final solution, but the authors are to be congratulated on furthering the art and demonstrating a useful technique capable of performing invaluable service in developing grinding wheels and grinding techniques that will be most suitable to critical applications.

It is hoped that this progress will be integrated into the overall progress of grinding in efficiency of metal removal and production of surface finish so that the proper balance is attained and due regard may be paid to the aspects developed in this work where residual stresses may play a significant role.

R. E. OGILVIE.² With regard to the x-ray stress-measuring technique described in the paper, the authors realize the insensitivity of the diffraction line used, but the writer questions whether they realize the exact magnitude of the errors. The stress constant is about 10,000 psi per 0.01 deg in theta. The writer is quite certain that the line position from a hardened-steel specimen cannot be determined better than 0.02 to 0.03 deg. This would mean a probable error in the measured stress of 20,000 to 30,000 psi.

The writer can see no reason why chromium radiation was not used in measuring the 211 line with the two-exposure technique. This would remove any doubt as to the effect on the interatomic spacing in the tempered and rehardened layer due to grinding. This effect has not been taken into account in this investigation. The stress constant would be about 1600 psi per 0.01 deg or an increase in sensitivity by a factor of 6.

The line position need not be the true position of the $K\alpha_1$ but any reference point may be used as long as it is consistent for the normal and oblique measurement. The peak position may be measured by fitting a parabola to the peak by the method of least squares or a technique used by Timken Roller Bearing Company where they extrapolate straight lines on the side of the peak, the intersection of these lines being the reference position. This procedure works in their case but it is believed in the higher-tempered steels investigated that this method will not work. The

¹ Bendix Products Division, South Bend, Ind.

² Massachusetts Institute of Technology, Cambridge, Mass.

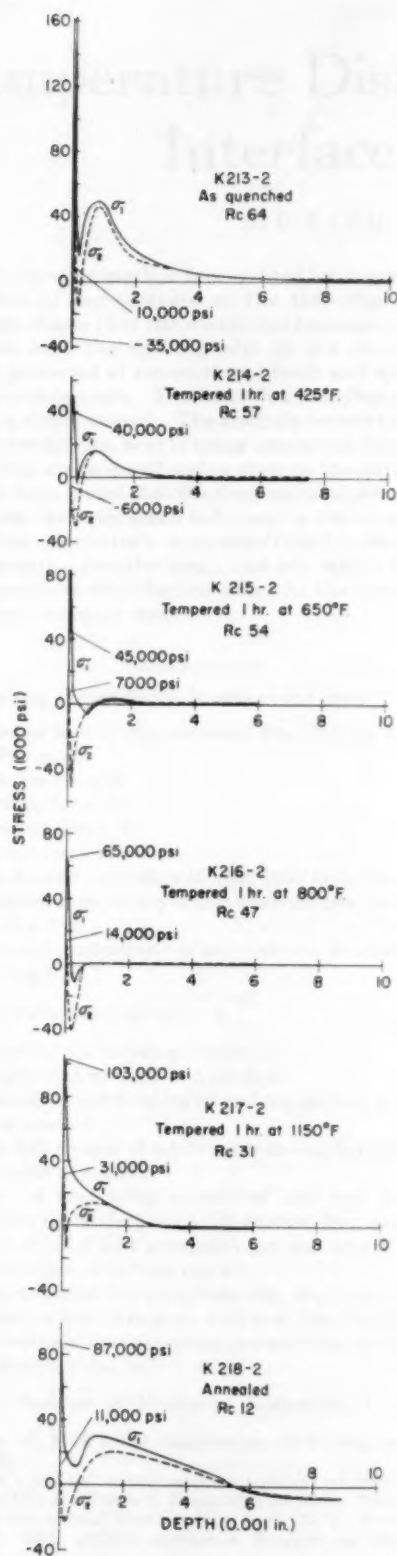


FIG. 11 EFFECT OF PRIOR HEAT-TREATMENT OF TOOL STEEL UPON RESIDUAL STRESSES INDUCED BY SURFACE GRINDING. PRINCIPAL STRESSES PARALLEL AND PERPENDICULAR TO THE DIRECTION OF ABRASIVE TRAVEL ARE DESIGNATED BY σ_1 AND σ_2 , RESPECTIVELY

center of gravity also may be used; however, this requires that the background be reached on both sides of the peak.

Considerable refinements are necessary but this is definitely a step in the right direction, and it is hoped that further work will be done along these lines.

H. R. LETNER.⁷ The results which the authors have presented are indeed a welcome addition to the relatively small fund of experimental information available on the effect of conditions of grinding and workpiece material upon residual grinding stresses. It is encouraging to see objective studies, such as this one, furnish information to supplant "grinding stresses are tensile stresses" generalities in engineering literature.

Although the authors' data adequately support their conclusions for the SAE 4340 steel which they used, our own research on another steel indicates that their fifth conclusion may be limited in its generality. Fig. 11 of this discussion typifies the variation of residual grinding stress with heat-treatment which we have consistently observed in experiments with manganese oil-hardening tool steel (0.92 C, 1.38 Mn, 0.019 P, 0.28 S, 0.013 Si, 0.45 Cr, 0.41 W). These data were obtained from 2-in-square specimens austenitized at 1475 F, quenched in oil, and tempered as indicated.

At the end of the grinding test in question, the specimens were 0.198 in. thick, 0.036 in. having been removed from the test surface of each specimen following heat-treatment. The six specimens were placed in a row on the chuck of a horizontal-spindle, reciprocating-table-type surface grinder and ground dry with a 7 $\frac{1}{2}$ -in. \times $\frac{1}{2}$ -in. \times $\frac{1}{4}$ -in. 38A461-J8VBE wheel. The wheel speed was 6000 fpm, traverse speed 60 fpm, unit crossfeed 0.050 in., and unit downfeed 0.001 in.

The average stress decreases with decreasing hardness until it reaches a minimum in the neighborhood of RC-54 to RC-47, then increases as the hardness continues downward. The thickness of the stressed layer exhibits a similar behavior, reaching a minimum value of 0.0007 in. in the RC-47 test piece, which is about one tenth of its value in the RC-64 and RC-12 specimens. Comparison of the behavior of this steel with the authors' SAE 4340 suggests that the effect of prior heat-treatment upon residual stresses induced by grinding is different for steels of different compositions.

AUTHORS' CLOSURE

The authors would like to thank the discussers for their comments. In regard to Mr. Ogilvie's comments it should be pointed out that the authors were well aware of the necessity of measuring accurately the position of the diffracted lines. It was for this reason that a least-squares fitting of the intensity data were carried out on a punched-card computer using a minimum of twenty-four points to obtain the curve. Repetitive traverses of the same surface resulted in maxima which were within ± 0.005 deg for 2θ . Chromium radiation would have been more satisfactory but at the time this work was being done this type of tube was unavailable. Admittedly, a two-exposure technique would have yielded more information but it is doubtful if the extra work involved could be justified.

We are grateful to Dr. Letner for presenting his results on tool steel. The fact that these data show a different sensitivity to work hardness is not surprising. Further work along this line can be expected to give rise to apparently anomalous results until the mechanism of producing residual stresses is better understood. The problem is not yet solved and future investigators would do well to be prepared to expect almost anything in the way of results. The problem is fascinating when one considers the potential for learning more about the mechanism of metal cutting itself.

⁷ Senior Fellow, Mellon Institute, Pittsburgh, Pa.

Temperature Distribution at the Tool-Chip Interface in Metal Cutting

By B. T. CHAO¹ AND K. J. TRIGGER,² URBANA, ILL.

A rapid, iterative method is presented for computing the distribution of temperature at the tool-chip interface. Calculation shows that the maximum temperature occurs at a point near the trailing edge of the contact when chips are produced at conventional feeds and speeds with sintered-carbide tools. The heat-flux distribution at the interface is also obtained. The analysis reveals that, under the usual conditions, heat is being conducted from the tool into the chip over a small region close to the cutting edge. It also has been found that the thermal conductivity of the tool material has only small influence on the mean cutting temperature. Relatively more significant is the effect on the temperature distribution. The role which the interface-temperature distribution plays in the formation of crater wear is pointed out.

NOMENCLATURE

The following nomenclature is used in the paper:³

- C = specific heat of chip material, Btu/(lb)(deg F)
- d = distance, in.
- F = friction force, lb
- F_c = cutting force, lb
- F_s = shearing force, lb
- F_t = thrust force, lb
- J = mechanical equivalent of heat, 9336 in-lb/Btu
- K_c = thermal conductivity of chip material, Btu/(min)(sq in.)(deg F/in.)
- K_t = thermal conductivity of tool material, Btu/(min)(sq in.)(deg F/in.)
- L_s = a dimensionless quantity = $\frac{v_f l_1}{2K_c}$
- $2l$ = length of the tool-chip contact, in.
- $2l_s$ = length of shear zone = $l_1 \csc \phi$, in.
- m = number of subdivisions on tool top surface along length of contact
- n = one half number of subdivisions on tool top surface along width of contact
- Q_i = rate of heat being transferred into tool top surface through lattice point at i th column, Btu/min
- q_i = rate of total heat generation per unit area of tool-chip interface, Btu/(min)(sq in.)
- $q_{i,c}$ = interface heat flux going into chip, Btu/(min)(sq in.)
- $q_{i,t}$ = interface heat flux going into tool, Btu/(min)(sq in.)
- q_s = rate of total heat generation per unit area of shear plane, Btu/(min)(sq in.)

- R_t = thermal number = $\frac{v_f l_1}{K_c}$
- r_t = chip-thickness ratio = t_1/t_2
- t_1 = thickness of "chip" before removal from workpiece (feed), in.
- t_2 = thickness of actual chip, in.
- V_c = cutting velocity, fpm
- v_c = cutting velocity, ipm
- v_f = chip-flow velocity, ipm
- v_s = velocity of shear, ipm
- w_1 = width of chip before removal from workpiece (depth of cut), in.
- w_2 = width of tool-chip contact, in.
- α = rake angle of tool
- ϵ = shearing strain undergone by chip during process of removal
- θ_0 = ambient temperature, or initial uniform temperature of workpiece, deg F
- $\bar{\theta}_s$ = average temperature of chip as it comes off shear zone, deg F
- $\Delta\bar{\theta}_s$ = average temperature rise of chip above ambient as it comes off shear zone = $\bar{\theta}_s - \theta_0$, deg F
- θ_i = local interface temperature, deg F
- $\Delta\theta_i$ = local interface temperature rise above ambient = $\theta_i - \theta_0$, deg F
- $\delta\theta_i$ = local interface temperature rise above $\bar{\theta}_s$ = $\theta_i - \bar{\theta}_s$, deg F
- κ_c = thermal diffusivity of chip material, sq in./min
- λ_s = fraction of shear-zone sensible heat conducted into workpiece
- ρ = weight density of work or chip material, pci
- ϕ = shear angle

INTRODUCTION

Tool-chip interface temperature has long been a subject of interest to many investigators in the field of metal cutting. Early in this century Taylor (1)⁴ reported that heat generated during the machining of metals played an important role in the wear of cutting tools. This instigated his search for a better tool material and led to the discovery of high-speed tool steels and their heat-treatment. Shore (2), Gottwein (3), and Herbert (4) were among the first to employ the tool-work thermocouple technique to determine quantitatively the average temperature at the tool-chip interface. Schwerd (5) determined the temperature field in the workpiece and the deformed chip material by the use of a radiation pyrometer. Schallbroch and Lang (6) employed thermosensitive colors to measure the surface temperature distribution of the tool and hence estimated the temperature at the cutting edge.

The first attempt to evaluate, analytically, the tool-chip interface temperature was made about 5 years ago (7). It was pointed out that in metal-machining operations there were two principal sources of heat generation, namely, the shear zone and the interface. Each contributes to the rise of temperature at the tool-chip

¹ Associate Professor of Mechanical Engineering, University of Illinois.

² Professor of Mechanical Engineering, University of Illinois. Mem. ASME.

³ Merchant's original nomenclature is retained wherever feasible. Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 23-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

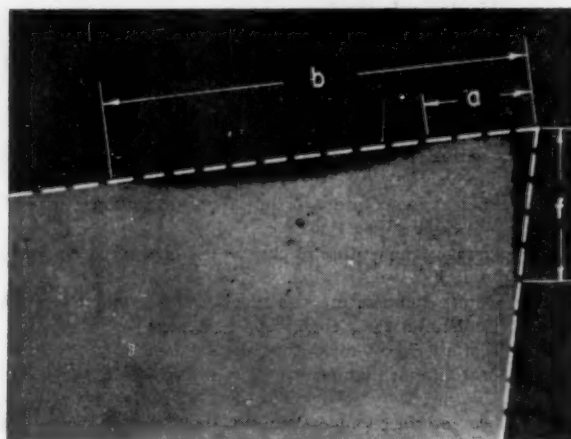
NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 16, 1954. Paper No. 54-A-115.

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

interface. More recently, Outwater and Shaw (8), Leone (9), Loewen and Shaw (10) introduced some refinements into the calculations. The partition of the shear-zone energy between the chip and workpiece was computed instead of being assumed, and thus the theoretically calculated average temperature was brought into closer agreement with the experimental data. However, the question of the temperature and heat-flux distribution over the tool-chip contact was not answered. This distribution is the primary objective of the present paper.

STUDIES ON MECHANISM OF TOOL WEAR ACCENTUATE NEED FOR INTERFACE-TEMPERATURE DISTRIBUTION

Experimental study of the development of top face wear of a cutting tool over wide ranges of cutting speeds and feeds and different tool-work combinations invariably reveals that cratering starts at some distance from the edge of the tool. Fig. 1 is a photographic record of the cross section of a carbide cutting tool after



Work material: AISI 4150 (modified) Q & T, 325 Bhn
 Tool material: Steel-cutting-grade carbide
 Tool shape: 0-6-7-7-10-0-0.015 in.
 Cutting speed: 367 fpm
 Feed: 0.00632 ipr
 Depth of cut: 0.125 in.
 Magnification: 145X
 Cutting time: 17.78 min

FIG. 1 CROSS SECTION OF A WORN TOOL

prolonged use. Such a profile is typical for continuous chip formation with either type 2 or type 3 chips. The lack of wear over a certain region behind the cutting edge cannot, therefore, be explained in terms of the built-up edge.

The fundamental mechanism of tool wear has been under study since 1952, in the Metal Cutting Research Laboratory at the University of Illinois and results have indicated that adhesion wear plays a major role in the formation of tool crater.¹ Under many circumstances, Holm's concept (11) of the probability of material removal, when interpreted in terms of aggregates of particles, can be used to advantage. For a given tool-work pair, it is thus deduced that the local temperature at the point of actual engagement is a principal factor. Conceivably, if this were the case, the crater formation can very possibly be the result of nonuniform temperature distribution at the interface. If, on the other hand, the nonuniformity of pressure distribution were the predominating factor, a wear profile such as illustrated in Fig. 1 could mean that either there would be a persistent built-up edge over the region *a* or the pressure there be negligible.

With the cutting conditions selected for a large number of

cases used in the experimental study, the chip showed little evidence of a pronounced built-up edge. It is difficult to comprehend the existence of an abrupt pressure drop near the cutting edge of the tool. If this cutting edge is dubbed slightly with a diamond hone the thrust force F_t is increased appreciably (some 10 to 15 per cent); in other words, the cutting edge is "in business." Nonuniform chip-pressure distribution is also at variance with Lee and Shaffer's analysis (12) of the chip-formation mechanics.

Following metallurgical reasoning and detailed examination of the wear process, Trent (13) concluded that cratering wear was essentially temperature-dependent. An estimate of the temperature distribution in the tool and at the interface was given but no calculation was attempted.

Inasmuch as the tool-chip interface is not accessible for study during chip formation, any experimental method to ascertain the distribution of temperature will of necessity be indirect. By first measuring the surface temperature of the tool, using thermosensitive colors, Bickel and Widmer (14) determined the interface-temperature distribution with the help of an electric-analog field plotter. For the one case cited the maximum temperature which occurred at the very edge of the tool was 727 C (1341 F) and the minimum was about 600 C (1112 F). The average cutting temperature as measured by the tool-work thermocouple technique was 560 C (1040 F) which is less than the minimum. Bickel and Widmer's result is at a variance with that of Trent. The desirability of a more accurate distribution of temperature at the interface is therefore evident.

DIFFICULTIES ARISING FROM USE OF BLOK'S PARTITION PRINCIPLE

One of the problems in the analytical evaluation of the metal-cutting temperature concerns the division of the interface heat as well as the division of heat generated at the shear zone. The conventional method for an approximate solution is the procedure first suggested by Blok (15). For the interface, it may be described briefly as follows: If the rate of heat generation per unit area of the interface is q_0 , and if λ_i designates the fraction of this interface heat being conducted into the tool, $(1 - \lambda_i)$ must represent the portion of the thermal energy being transferred into the chip provided that all heat losses are ignored. An expression is then derived which gives the average temperature over the tool-chip contact due to a stationary heat source of intensity $\lambda_i q_0$ on the top surface of the tool. A second expression which again gives the average temperature over the same contact can be formulated by considering the moving heat source of intensity $(1 - \lambda_i) q_0$ traveling with the chip-flow velocity at the separating surface of the chip. By equating these two expressions, the unknown λ_i can be evaluated. A similar method has been used for the determination of the division of heat at the shear zone.

While Blok's procedure gives a convenient means of estimating the division of heat at moving contacts, it oftentimes results in an anomaly when the distribution of temperature is sought. This is due to the rather unrealistic assumption used in the procedure, namely, that the distribution of heat flux over the contact region is uniform. This latter viewpoint can probably be best illustrated by analyzing the cutting data which follow:

Work material: NE9445 mill-annealed, 183 Bhn
 Tool material: Triple carbide
 Tool shape: 0-4-7-7-6-0-0-1/4
 Depth of cut: $w_1 = 0.102$ in.
 Feed: $f_1 = 0.0098$ ipr
 Cutting speed: $V_s = 300$ fpm
 Tool forces: $F_s = 378$ lb, $F_t = 192$ lb
 Chip-thickness ratio: $r_1 = 0.375$
 Length of tool-chip contact: $2l = 0.0476$ in.

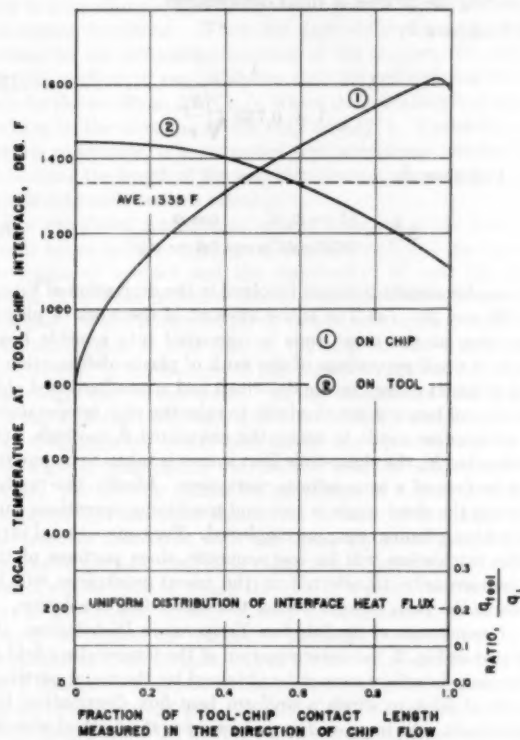
¹Details will be reported in a subsequent paper.

Width of tool-chip contact: $w_2 = 0.119$ in.

Room temperature: 65 F

Interface temperature: 1300 F, as determined by tool-work thermocouple

Pursuant to the procedure as outlined in a previous paper (7), the average temperature of the chip as it leaves the shear zone has been found to be 748 F. With a uniformly distributed heat flux at the chip surface, the temperature distribution on the chip surface can be calculated readily from Jaeger's method of analysis (16). The result is shown by curve 1 in Fig. 2. It is seen that the temperature distribution is not uniform as a result of the chip motion; rather, the maximum shifts toward the separation point of the chip and the tool as speed and/or feed are increased.



- (1) Calculated from point of view of chip
(2) Calculated from point of view of tool

FIG. 2 INTERFACE-TEMPERATURE DISTRIBUTION WITH UNIFORMLY DISTRIBUTED HEAT FLUX

With an assumed uniform heat flux at the chip surface, the heat flux at the tool face over the contact likewise must be uniform. If heat loss is ignored, the resulting temperature distribution at the common interface calculated from the point of view of conduction in the tool is given by curve 2, Fig. 2. Use is made of Blok's partition principle in the computation. While the calculated and the measured average temperature are in good agreement, the anomaly of having two completely different temperature distributions at the common interface is evident. This has prompted the authors to re-examine the conventional assumptions and to introduce an improved method of calculation. It should be remarked that the difficulty encountered in the calculation of the true temperature field at the common interface was recognized by Blok some 15 years ago. However, no satisfactory solution had been found.

AN APPROXIMATE BUT IMPROVED PROCEDURE

From the preceding analysis it is concluded that, in order to bring the two temperature-distribution curves near to coincidence, the heat-flux distribution necessarily must be nonuniform. Obviously, the heat flow into the tool at, and in the neighborhood of, the cutting edge should be below average or even negative. In the vicinity of the tool-chip separation point, it should be considerably above average. Such a distribution is depicted schematically by curve *b-b* in Fig. 3.

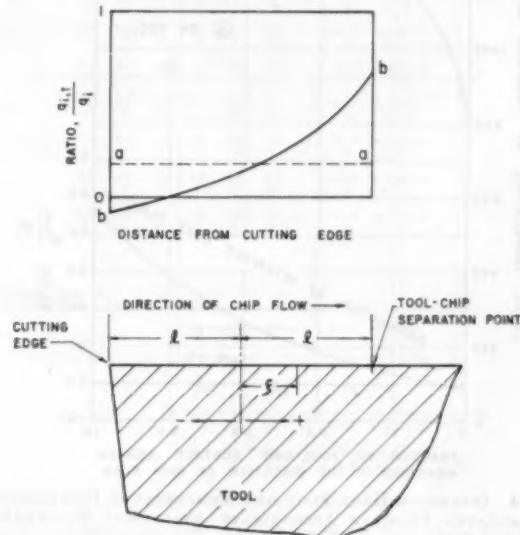


FIG. 3 NONUNIFORM INTERFACE HEAT-FLUX DISTRIBUTION

The first attempt made to elucidate the temperature field at the interface was an approximate analytical procedure in which a heat-flux distribution of suitable form was assumed. One of the several distributions used is an exponential relationship given by

$$\frac{q_{i,s}}{q_i} = C_1 - C_2 e^{-C_3 \left(1 - \frac{\xi}{l}\right)} \quad [1]$$

in which

$q_{i,s}$ = rate of heat transferring into chip per unit area at location ξ

q_i = rate of total interface heat per unit area of contact (it is assumed to be uniform)

$2l$ = length of tool-chip contact

ξ = distance measured from mid-length of contact, positive in direction of chip flow

C_1, C_2, C_3 = constants, to be selected such that condition of "compatibility" can be fulfilled approximately. The latter requires that the calculated temperature distribution at the interface due to the moving heat source at the separating surface of the chip and that due to the stationary heat source at the top surface of the tool are identical. The determination of the constants C_1, C_2 , and C_3 involves a cut-and-try procedure.

Since a better, iterative method has been developed as a consequence of this study, only results of the approximate procedure are given here. Cutting conditions and other relevant experimental data required in the computation are the same as those used in the preparation of Fig. 2.

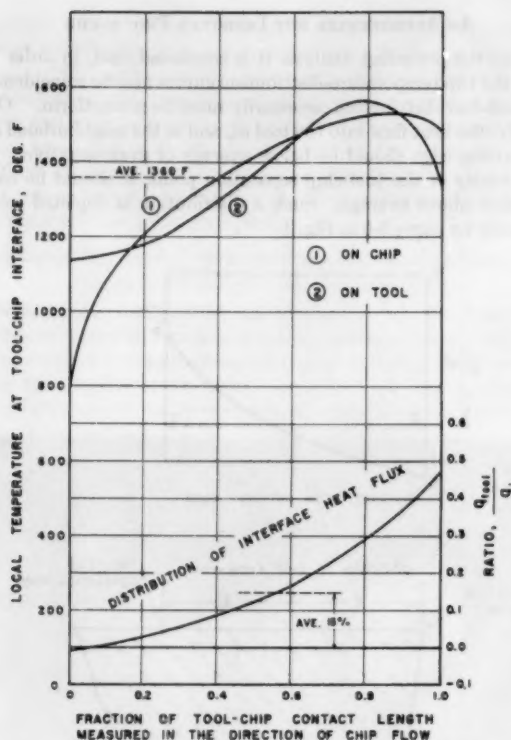


FIG. 4 INTERFACE HEAT FLUX AND TEMPERATURE DISTRIBUTION CALCULATED FROM AN APPROXIMATE ANALYTICAL PROCEDURE

With $C_1 = 1.08$, $C_2 = 0.55$, and $C_3 = 1.00$, the heat-flux distribution is given by the lower curve in Fig. 4. The calculated temperature field on the chip and on the tool are shown, respectively, by curves 1 and 2 of the same figure. The true common local interface temperature can be approximated by finding the average of the two distributions. It is seen that the maximum temperature does not occur at the mid-length of the contact but is closer to the separating edge of the chip and the tool. If the arithmetic mean is calculated, it has been found to be 1360 F which compares favorably with the average value shown in Fig. 2. While a more realistic interface-temperature distribution has been achieved, the method has the inherent drawback of using the cut-and-try procedure in finding the three constants in Equation [1]. It is tedious and time-consuming. A heat-flux distribution expressible in the form of second-degree polynomials has been used, but the results are even less satisfactory.

A RAPID, ITERATIVE PROCEDURE FOR FINDING TEMPERATURE DISTRIBUTION AT TOOL-CHIP INTERFACE

Since the problem of finding the temperature field over the tool-chip contact is of great interest in the study of the basic mechanism of crater wear, a method of solution which has considerable merit over that previously given is now described in detail. It may be pertinent to point out that the method is quite general in application and can be used readily in ascertaining the distribution of temperature at sliding contacts.

As pointed out previously, the tool-chip interface temperature is made up of two components: (a) that due to chip deformation at the shear zone, and (b) that due to sliding of the heated chip on the tool surface. For completeness, the equations required for the computation of the average shear-plane temperature are reproduced in the following. Reference is made to Bibliography items (7) and (10) for detailed discussion.

I Recommended Procedure for Calculation of Average Shear-Plane Temperature:

(i) Find ϕ , ϵ , and F_s from the following equations

$$\tan \phi = \frac{r_f \cos \alpha}{1 - r_f \sin \alpha} \quad [2]$$

$$\epsilon = \cot \phi + \tan (\phi - \alpha) \quad [3]$$

and

$$F_s = F_s \cos \phi - F_t \sin \phi \quad [4]$$

(ii) Compute R_s . κ_s is first estimated and corrected later when $\bar{\theta}_s$ is known. It is the mean thermal diffusivity of the work material during the process of shear deformation.

(iii) Compute λ_s

$$\lambda_s = \frac{1}{1 + 0.753 \sqrt{\frac{R_s}{\epsilon}}} \quad [5]^*$$

(iv) Calculate $\bar{\theta}_s$

$$\Delta \bar{\theta}_s = \frac{(1 - \lambda_s) F_s}{9336 w t_c \rho \cos (\phi - \alpha)} \quad [6]^*$$

Two major assumptions are involved in the derivation of Equations [5] and [6]: (a) The entire amount of the work of plastic deformation at the shear zone is converted into sensible heat. Actually a small percentage of the work of plastic deformation is stored as latent energy in the deformed and strain-hardened chip material, and hence is not available to raise the chip temperature. This assumption tends to make the calculated $\bar{\theta}_s$ too high. (b) In evaluating λ_s , the shear-zone heat source is taken to be moving on the surface of a semi-infinite workpiece. Ideally, this applies only when the shear angle is zero and machining operations such as broaching, planing, etc., are employed. For conventional turning, the calculation will be less accurate, since portions of the heat momentarily transferred to the uncut workpiece will be present in the chip formed during the succeeding revolution.

II *Computation of the Interface Temperature Distribution.* As illustrated in Fig. 2, the determination of the temperature field at the common interface cannot be obtained by the usual partition principle of Blok in which a uniform heat-flux distribution has been assumed. An improved procedure using the method of iteration is presented. Inherent in the procedure is the further advantage that both the temperature and heat-flux distribution are arrived at simultaneously.

For the first approximation, all the heat liberated at the tool-chip interface is considered to be transferred into the chip. Hence at this stage of computation the heat is taken to be uniformly distributed over the contact, and Jaeger's moving-source expression can be applied directly to calculate the local interface temperature. Thus the temperature rise above $\bar{\theta}_s$ at a point distant ξ from the mid-length of contact (see Fig. 2 for sign convention) is

$$\delta \theta_i = \frac{2q_i}{\pi C \rho v_f} \int_{E-L}^{E+L} e^{-u K_0(|u|)} du \quad [7]^*$$

where

$$E = \frac{v_f \xi}{2\kappa_s}, \text{ and } L = \frac{v_f l}{2\kappa_s}$$

* See Appendix for derivation.

† See Bibliography (16).

and K_0 is the modified Bessel function of second kind and zero order. The assumptions involved are (a) the chip can be viewed as a semi-infinite solid with a band source of width $2l$ and intensity q_i traveling at the velocity v_f over its separating surface, (b) constant thermal properties of the chip material at temperature $\bar{\theta}_0$ can be used, (c) a quasi-stationary temperature distribution has been reached, and (d) the heat losses are negligible. At cutting speeds and feeds commonly employed with cemented-carbide tools, these assumptions have been found, in general, to be reliable.*

It is pertinent to point out that in so far as the calculation of the temperature distribution on the chip surface is concerned, Jaeger's analytical procedure is more convenient than the usual numerical method of analysis. This is true even if the heat flux at the interface is nonuniform. The latter case arises in the second and all succeeding iterations. When the finite-difference equivalent is written for the differential equation of the temperature field in a moving medium, it can be shown that the convergence criterion calls for the condition $\Delta x \leq \kappa_c/v_f$ where Δx is the length of the subdivision in the direction of the chip flow (17). Under the usual cutting conditions, it is estimated that nearly one hundred divisions along the length of the tool-chip contact are necessary. This entails extremely tedious calculation.

The calculated temperature profile obtained in the first iteration is taken to be coexisting on the top surface of the tool over the region of contact and the distribution of heat flux determined. The latter is then deducted from the total, uniform interface heat. Their difference gives, as a second approximation, the new distribution of heat flux on the chip surface which is now nonuniform. A more accurate interface-temperature distribution accordingly can be calculated. It has been found that three or four iterations are sufficient to achieve suitable accuracy.

To determine the steady heat flux at the top surface of the tool with a given temperature distribution over the area of tool-chip contact, the familiar relaxation procedure of Emmons was first attempted but was later abandoned because of the practical inconveniences associated with the unusually small proportion of the area of contact relative to other tool dimensions. A new approach was evolved in which the idea of a point heat source was used to advantage. Briefly, it may be summarized as follows:

As in the case of the relaxation methods, the solid (tool) is not considered as a continuum, but as a network of conducting rods. Fig. 5 illustrates diagrammatically the rectangular contact at the top surface of a wide tool. A cubic lattice is shown with each side of length Δx . On the top surface the location of each lattice point is designated by two numerals referring, respectively, to the number of the column and the row. A cross section of the tool through OX , which in the present case is also an axis of symmetry, is shown in the same figure. The lattice points $1', 2', \dots$ etc., in that section are at a distance Δx below the top surface.

In the discussion which follows, both the rake and the clearance angles are assumed to be zero. Preliminary calculation reveals that the heat loss from the tool surface is relatively unimportant and is thus ignored. Heat flux into the tool varies along rows of the lattice points but is considered uniform along a given column. Hence, referring to Fig. 5, $Q_{1,0} = Q_{1,1} = Q_{1,2} = \dots = Q_{1,n}$ and $Q_{2,0} = Q_{2,1} = Q_{2,2} = \dots = Q_{2,n}$; etc.

It is known that the steady temperature rise at any point P in an infinite solid due to a point heat source continuously liberating heat at the rate of Q per unit time is $Q/(4\pi Kd)$ where K is the thermal conductivity of the solid material and d is the distance between the source and the point P (18). For a semi-infinite solid with the point source lying in the plane surface which is otherwise perfectly insulated, the corresponding rise in temperature is

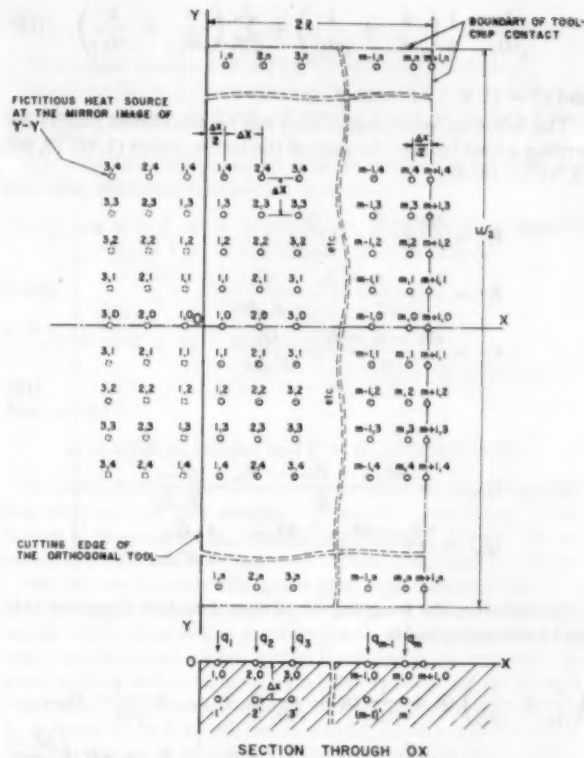


FIG. 5 NETWORK ON TOP SURFACE OF WIDE TOOL OVER ITS RECTANGULAR TOOL-CHIP CONTACT

$Q/(2\pi Kd)$. Accordingly, in Fig. 5 the temperature rise at $1', 2', \dots$ etc. can be computed readily by imagining the tool to be extended to the half space left of YY with fictitious heat sources situated at the mirror image of the actual sources. Each fictitious source has a strength identical to that of its mirror image. Hence, if Q_i represents the rate of heat flowing into the tool over the area $(\Delta x)^2$ at the i th column, $1'd_{i,j}$ and $1'd_{i,j}$ denote, respectively, the distance from the lattice point (i, j) and its image (i, j) to point $1'$, the steady temperature rise above θ_0 at $1'$ is

$$\Delta\theta_{1'} = \theta_{1'} - \theta_0 = \frac{1}{2\pi K} \sum_{i=1}^m Q_i \left(\frac{1}{1'd_{i,0}} + \frac{1}{1'd_{i,0}} \right) + \frac{1}{\pi K} \sum_{i=1}^m \sum_{j=1}^n Q_i \left(\frac{1}{1'd_{i,j}} + \frac{1}{1'd_{i,j}} \right) \quad [8]$$

Other symbols are as listed in the Nomenclature. If one defines

$$\frac{1}{1'D_i} = \frac{1}{2} \left(\frac{1}{1'd_{i,0}} + \frac{1}{1'd_{i,0}} \right) + \sum_{j=1}^n \left(\frac{1}{1'd_{i,j}} + \frac{1}{1'd_{i,j}} \right) \quad [9]$$

Equation [8] becomes

$$\Delta\theta_{1'} = \theta_{1'} - \theta_0 = \frac{1}{\pi K} \sum_{i=1}^m \frac{Q_i}{1'D_i} \quad [10]$$

In general

$$\Delta\theta_{r'} = \theta_{r'} - \theta_0 = \frac{1}{\pi K} \sum_{i=1}^m \frac{Q_i}{r'D_i} \quad [11]$$

with

*See Bibliography (7).

$$\frac{1}{r'D_i} = \frac{1}{2} \left(\frac{1}{r'd_{i,0}} + \frac{1}{r'd_{i,0}} \right) + \sum_{j=1}^n \left(\frac{1}{r'd_{i,j}} + \frac{1}{r'd_{i,j}} \right) \dots [12]$$

and $r' = 1', 2', \dots, m'$.

The following relationships also can be established readily by writing a heat balance for each of the lattice points (1, 0), (2, 0), (3, 0) ... (m, 0)

$$\left. \begin{aligned} \theta_1' &= \frac{3\theta_1 - \theta_2}{2} - \frac{Q_1}{K_i \Delta x} \\ \theta_2' &= \frac{4\theta_2 - \theta_1 - \theta_3}{2} - \frac{Q_2}{K_i \Delta x} \\ \theta_3' &= \frac{4\theta_3 - \theta_2 - \theta_4}{2} - \frac{Q_3}{K_i \Delta x} \\ &\vdots \\ Q_{(m-1)'} &= \frac{4\theta_{m-1} - \theta_{m-2} - \theta_m}{2} - \frac{Q_{m-1}}{K_i \Delta x} \\ Q_{m'} &= \frac{9\theta_m - 2\theta_{m-1} - 4\theta_{m+1}}{3} - \frac{4}{3} \frac{Q_m}{K_i \Delta x} \end{aligned} \right\} \dots [13]$$

Substituting the foregoing set of equations into Equation [11] and rearranging yields

$$\left. \begin{aligned} \left(\frac{\pi}{\Delta x} + \frac{1}{1'D_1} \right) Q_1 + \frac{1}{1'D_2} Q_2 + \frac{1}{1'D_3} Q_3 + \dots + \frac{1}{1'D_{m-1}} Q_{m-1} \\ + \frac{1}{1'D_m} Q_m = \pi K_i \left(\frac{3\theta_1 - \theta_2}{2} - \theta_0 \right) \\ \frac{1}{2'D_1} Q_1 + \left(\frac{\pi}{\Delta x} + \frac{1}{2'D_2} \right) Q_2 + \frac{1}{2'D_3} Q_3 + \dots + \frac{1}{2'D_{m-1}} Q_{m-1} \\ + \frac{1}{2'D_m} Q_m = \pi K_i \left(\frac{4\theta_2 - \theta_1 - \theta_3}{2} - \theta_0 \right) \\ \vdots \\ \frac{1}{(m-1)'D_1} Q_1 + \frac{1}{(m-1)'D_2} Q_2 + \frac{1}{(m-1)'D_3} Q_3 + \dots + \left(\frac{\pi}{\Delta x} + \frac{1}{(m-1)'D_{m-1}} \right) Q_{m-1} \\ + \frac{1}{(m-1)'D_m} Q_m = \pi K_i \left(\frac{4\theta_{m-1} - \theta_{m-2} - \theta_m}{2} - \theta_0 \right) \\ \frac{1}{m'D_1} Q_1 + \frac{1}{m'D_2} Q_2 + \frac{1}{m'D_3} Q_3 + \dots + \frac{1}{m'D_{m-1}} Q_{m-1} \\ + \left(\frac{4}{3} \frac{\pi}{\Delta x} + \frac{1}{m'D_m} \right) Q_m = \pi K_i \left(\frac{9\theta_m - 2\theta_{m-1} - 4\theta_{m+1}}{3} - \theta_0 \right) \end{aligned} \right\} \dots [14]$$

We thus have m simultaneous linear algebraic equations to be solved for m unknowns, namely, Q_1, Q_2, \dots, Q_m . It may be remarked that the coefficients of the first and the last equation have a slightly different pattern when compared with those of others. This is due to the fact that the first column of the lattice points is at a distance $\Delta x/2$ from the edge YY . Likewise, the m th column is also distant $\Delta x/2$ from the edge where tool-chip separation starts. As common with other methods of finite divisions, Equations [8] to [14], inclusive, are not exact. The larger the number of subdivisions, the better will be the accuracy. In many practical cases of local interface-temperature calculation, it has been found that 10 or 20 subdivisions would be adequate. Standard programming is available for solving simultaneous

linear algebraic equations by the digital computer, the "Illiac."

While the previous equations are derived under the condition that cutting is being done at the edge of a wide tool (orthogonal cutting), the method of calculation and the principle of image source can be extended readily to the case in which cutting is taking place at the corner of a rectangular tool (conventional turning). Fig. 6 illustrates the distribution of the real and fictitious heat sources in this latter case. Sources labeled with the same numeral are identical in strength. The procedure for either edge or corner cutting is given in the following illustration.

Sample Computation. Cutting and other relevant data are as follows:

Work material: AISI 4150 (mod) annealed, 197 Bhn

Tool material: Steel cutting grade carbide

Tool shape: 0-6-7-7-10-0-0.015 in., ($\alpha = 6^\circ$)

Depth of cut: $w_1 = 0.125$ in.

Feed: $t_1 = 0.00632$ ipr

Cutting speed: $V_c = 456$ fpm

Tool forces: $F_s = 280$ lb, $F_t = 135$ lb

Chip-thickness ratio: $r_1 = 0.397$

Length of tool-chip contact: $2l = 0.0268$ in.

Width of tool-chip contact: $w_2 = 0.137$ in.

Room temperature: $\theta_0 = 78$ F

Interface temperature: 1385 F (average of 5 readings), as determined by tool-work thermocouple

I Calculation of θ_s :

$$(i) \tan \phi = \frac{0.397 \cos 6^\circ}{1 - 0.397 \sin 6^\circ} = 0.412, \phi = 22.4^\circ$$

$$\epsilon = \cot 22.4^\circ + \tan (22.4^\circ - 6^\circ) = 2.72$$

$$F_s = 280 \cos 22.4^\circ - 135 \sin 22.4^\circ = 207.2 \text{ lb}$$

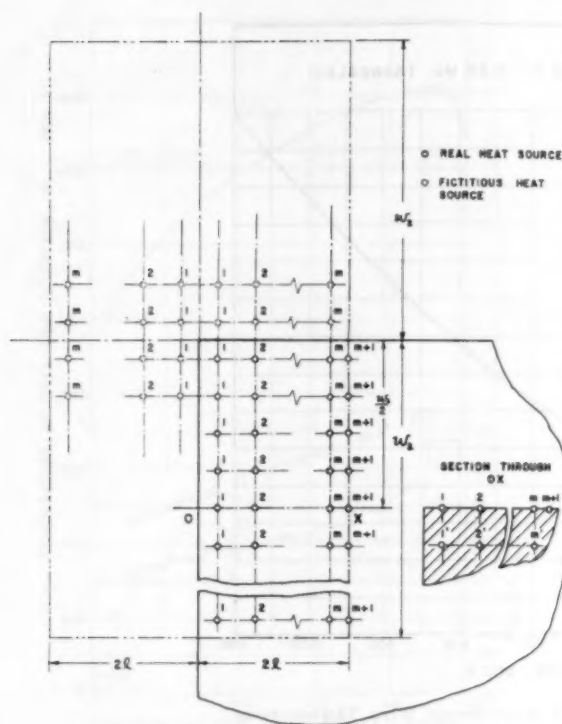
(ii) Estimate the value of θ_s , e.g., 650 F, and hence

$$\frac{\theta_s + \theta_0}{2} = 364 \text{ F}$$

From Fig. 7, κ_s at 364 F is 0.966 sq in/min

$$\therefore R_i = \frac{456 \times 12 \times 0.00632}{0.966} = 34.6$$

* The "Illiac," which is the digital computing machine at the University of Illinois, can handle such equations up to $m = 39$.



SOURCES LABELED WITH SAME NUMERAL HAVE IDENTICAL INTENSITY
 FIG. 6 REAL AND FICTITIOUS HEAT SOURCES FOR TEMPERATURE CALCULATION OF A CORNER-CUTTING RECTANGULAR TOOL

and

$$\bar{\theta}_s = 609 + 78 = 687 \text{ F}$$

The weight density ρ of AISI 4150 steel is 0.283 lb/in.³

(v) Steps (ii) to (iv) are repeated if original estimate of $\bar{\theta}_s$ is in considerable error. Refined calculation gives $\Delta\bar{\theta}_s = 606 \text{ F}$ and $\bar{\theta}_s = 684 \text{ F}$.

(II) For the computation of local interface temperature, the following additional information is required

$$F = F_1 \cos \alpha + F_2 \sin \alpha = 135 \cos 6^\circ + 280 \sin 6^\circ = 163.5 \text{ lb}$$

$$V_f = V_s r_1 = 181.2 \text{ fpm}, \quad v_f = 2174 \text{ ipm}$$

Hence

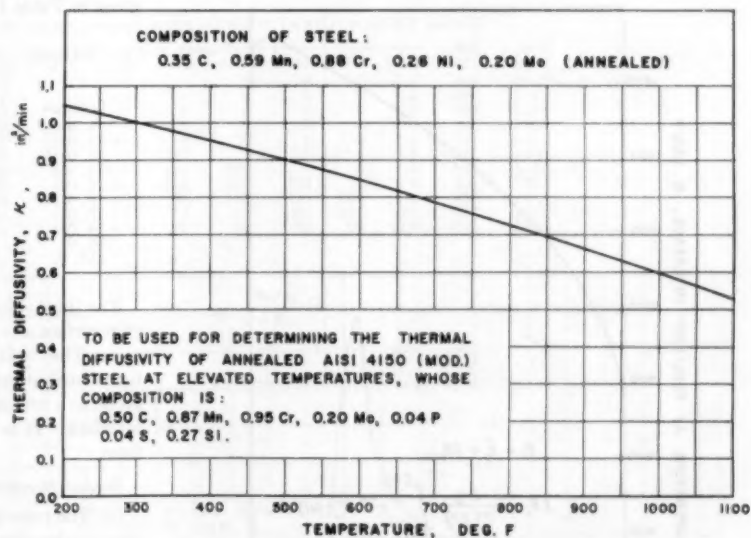
$$q_s = \frac{F v_f}{9336 \times 2l \times w_2} = \frac{163.5 \times 2174}{9336 \times 0.0268 \times 0.137} = 10,350 \text{ Btu}/(\text{min})(\text{sq in.})$$

Also, at 684 F,

$$\kappa_s = 0.796 \text{ sq in}/\text{min}, \text{ and } C = 0.140 \text{ Btu}/(\text{lb})(\text{F})$$

Available data on the thermal conductivity K_i of carbide cutting materials are very meager. Information concerning the influence of temperature is virtually nonexistent. Dawahl (19) reported the values of 0.19 and 0.09 cal/(sec)(cm²)(deg C/cm) for a straight tungsten-carbide grade and a steel-cutting grade of carbide, respectively. However, it is known that their thermal conductivity depends to a great extent on the chemical composition and impurities. Even within the category of steel cutting-grade carbide, values ranging from 0.068 to 0.136 cal/(sec)(cm²)(deg C/cm) have been reported (20). For the present calculation, K_i is taken to be 0.10 cal/(sec)(cm²)(deg C/cm) or 0.034 Btu/(min)(sq in.)(deg F/in.) which is the value used in reference (7). A higher value has been used in reference (10).

FIG. 7 VARIATION OF THERMAL DIFFUSIVITY OF AN ALLOY STEEL WITH TEMPERATURE



$$(iii) \lambda_s = \frac{1}{1 + 0.753 \sqrt{\frac{34.6}{2.72}}} = 0.271$$

(iv) At 364 F, $C = 0.123 \text{ Btu}/(\text{lb})(\text{F})$ (see Fig. 8) and thus

$$\Delta\bar{\theta}_s = \frac{(1 - 0.271) \times 207.2}{9336 \times 0.125 \times 0.00632 \times 0.123 \times 0.283} \times \frac{0.9945}{0.9593} = 609 \text{ F}$$

First Iteration:

(i) With all the interface heat transferred into the chip, the temperature distribution over the area of contact can be readily computed from Equation [7]. The result is shown graphically in Fig. 9.

(ii) For convenience, select $m = 10$, thus $\Delta x = 0.00268 \text{ in.}$ and $n = 25$. Compute the various distances $r'd_{i,j}$ and $r'd_{1,j}$ and thence $1/r'd_i$ by using Equation [12].

(iii) Calculate all the coefficients in the set of Equations [14]. $\theta_1, \theta_2 \dots \theta_{m+1}$ are taken from Fig. 9.

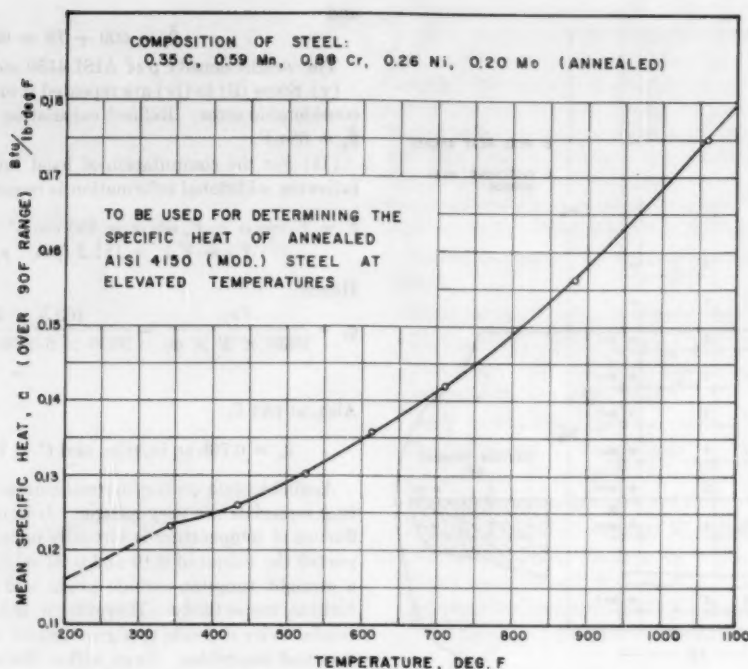


FIG. 8 VARIATION OF SPECIFIC HEAT OF AN ALLOY STEEL WITH TEMPERATURE

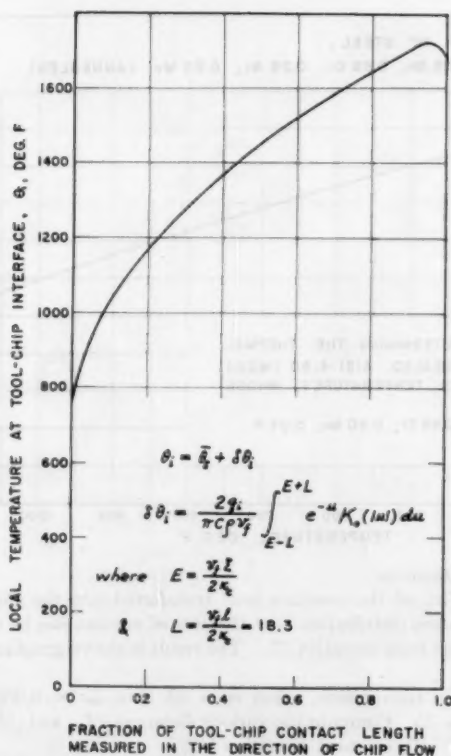


FIG. 9 TEMPERATURE DISTRIBUTION ALONG TOOL-CHIP CONTACT WHEN TOOL MATERIAL IS A PERFECT INSULATOR

(iv) Solve for $Q_1, Q_2 \dots Q_m$. Compute $q_{i,c}$. The results are given in Table 1.

TABLE 1 HEAT-FLUX DISTRIBUTION AT TOOL-CHIP INTERFACE

Location	Q_i Btu/min	$q_{i,c}$ Btu/(min)(sq in.)	Ratio, $q_{i,c}/q_i$
(1, 0)	-19.06×10^{-3}	13000	-0.256
(2, 0)	2.23	10040	0.030
(3, 0)	4.39	9740	0.059
(4, 0)	6.32	9470	0.085
(5, 0)	7.38	9320	0.099
(6, 0)	9.12	9080	0.123
(7, 0)	11.20	8790	0.151
(8, 0)	13.84	8420	0.185
(9, 0)	18.76	7740	0.252
(10, 0)	27.88	6470	0.373

The third column of Table 1 lists new values of heat flux on the chip surface at various locations along the length of tool-chip contact. They will be used to compute the new interface-temperature distribution in the 2nd iteration. The lower curve in Fig. 10 gives, as a first approximation, the distribution of heat flux at the interface. It is seen that this distribution is extremely nonuniform.

Second Iteration:

(i) The principle of superposition is now employed to calculate the new, local interface temperature. For each lattice point, the heat "load" is uniform over a span of length Δx . Hence Equation [7] is applicable provided that q_i is replaced by $q_{i,c}$ and $2l$ by Δx . Since, in the present calculation, $m = 10$, the corresponding value of

$$L \left(= \frac{v_i \Delta x}{4\kappa_s} \right) = 1.83$$

The temperature distribution plotted in dimensionless co-ordinates due to any such heat load is shown in Fig. 11. The origin is taken at the mid-length of the span Δx .

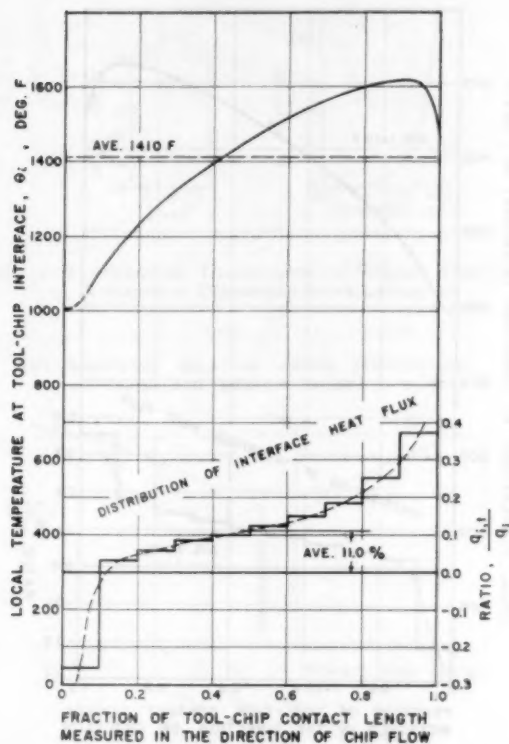


FIG. 10 INTERFACE HEAT FLUX AND TEMPERATURE DISTRIBUTION (EDGE CUTTING)—FIRST APPROXIMATION

(ii) Construct a new temperature-distribution curve on the chip surface over the entire length of the tool-chip contact. Use is made of Fig. 11. As an illustration, consider the location (3, 0). Temperature rise due to heat flux at (1, 0) where $q_{i,c} = 13,000$ Btu/(min)(sq in.) is

$$\frac{2q_{i,c}}{\pi C p v_f} \times \text{ordinates of curve in Fig. 11 at } \frac{\xi}{\Delta x/2} = 4$$

or

$$\frac{2 \times 13,000}{\pi \times 0.140 \times 0.283 \times 2174} \times 1.69 = 162 \text{ deg F}$$

Similarly, the temperature rise due to heat flux at (2, 0) with $q_{i,c} = 10,040$; (3, 0) with $q_{i,c} = 9,740$ and (4, 0) with $q_{i,c} = 9,470$ are 178, 258, and 1 deg F, respectively. The results are summarized in Table 2.

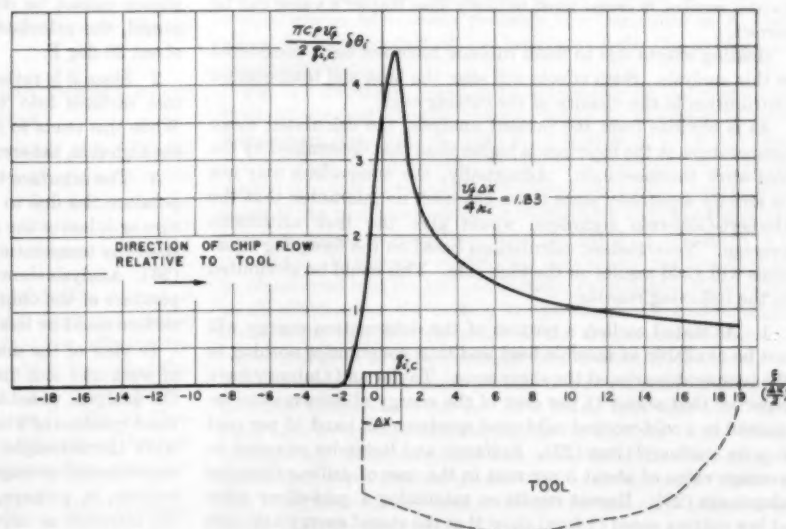
The upper curve in Fig. 10 shows, graphically, the new temperature distribution at the tool-chip interface. The arithmetic average is 1410 F.

The foregoing method of calculation can be repeated if better accuracy is desired. However, in the present case it has been found that iteration beyond the second is not necessary in view of the uncertainties involved in the various assumptions used. This is illustrated by Fig. 12 in which the computations have been carried out to the fourth iteration. It is noted that there is little change in the temperature and heat-flux distribution when compared with Fig. 10.

When cutting at the corner of a large tool, the calculated interface temperature will be slightly higher although comparatively less heat is conducted into the tool. Fig. 13 illustrates the distribution of temperature and heat flux at the tool-chip interface

TABLE 2 LOCAL TEMPERATURE RISE DUE TO VARIOUS HEAT LOADS

Location (see Fig. 5)	13,000	10,040	9,740	9,470	9,320	9,080	8,790	8,420	7,740	6,470	θ_i deg F	θ_i deg F
(1, 0).....	344	2	0	0	0	0	0	0	0	0	346	1030
(2, 0).....	230	266	1	0	0	0	0	0	0	0	496	1181
(3, 0).....	162	178	258	1	0	0	0	0	0	0	599	1283
(4, 0).....	131	125	173	251	1	0	0	0	0	0	681	1365
(5, 0).....	113	101	121	168	247	1	0	0	0	0	751	1435
(6, 0).....	101	87	98	118	165	241	1	0	0	0	811	1495
(7, 0).....	94	78	85	95	116	161	233	1	0	0	863	1547
(8, 0).....	86	72	76	82	94	113	156	223	1	0	903	1587
(9, 0).....	81	67	70	74	81	91	110	149	205	1	929	1613
(10, 0).....	77	63	65	68	73	79	89	105	137	171	927	1611
(11, 0).....	75	59	61	63	67	71	77	85	96	115	769	1453

FIG. 11 SURFACE TEMPERATURE DISTRIBUTION OF A SEMI-INFINITE SOLID CAUSED BY SLIDING FRICTION AT CONTACT OF WIDTH Δx 

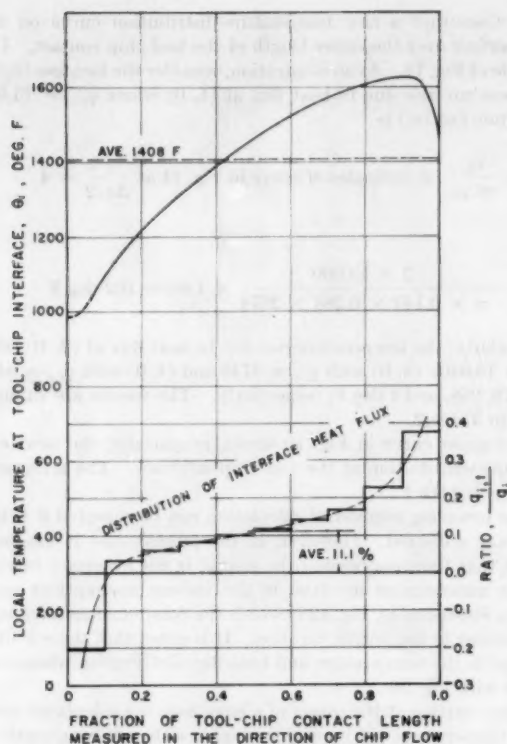


FIG. 12 TOOL-CHIP INTERFACE HEAT FLUX AND TEMPERATURE DISTRIBUTION (EDGE CUTTING)—FINAL RESULT

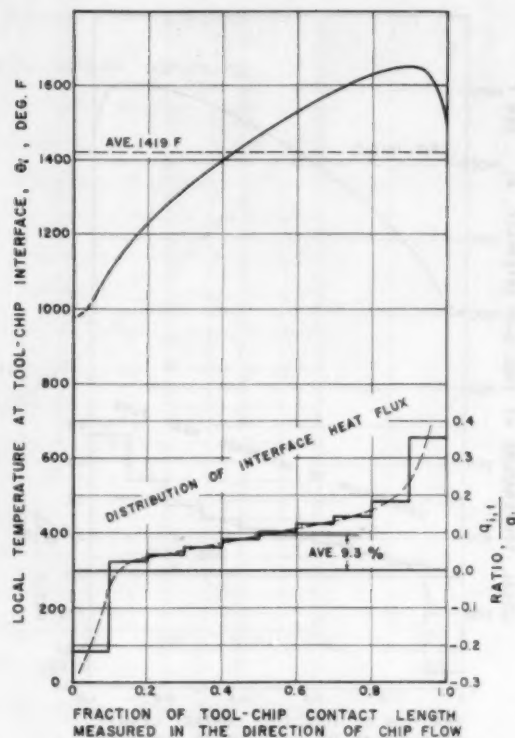


FIG. 13 TOOL-CHIP INTERFACE HEAT FLUX AND TEMPERATURE DISTRIBUTION (CORNER CUTTING)—FINAL RESULT

under this condition. As for the case of edge-cutting, the variation of temperature in a direction normal to chip flow is not considered. In either instance it is seen that heat is actually flowing from the tool into the chip over a small distance near the cutting edge. In a communication to Trent, Rapier (13) claimed a similar finding but neither data nor the method of calculation were given. Under a cutting condition which did not differ radically from that used in the present illustration, a maximum temperature of 1300 C (2372 F) was reported. In the light of experimental results, it seems most unlikely that Rapier's value can be correct.

Heating effects due to flank rubbing have not been considered in this analysis. Such effects will alter the heat and temperature distribution in the vicinity of the cutting edge.

As is obvious from the present analysis, the calculated mean temperature at the interface is higher than that determined by the tool-work thermocouple. Admittedly, the comparison may not be strictly equitable, since there has been no guarantee that the Herbert-Gottwein technique would give the true arithmetic average. Nevertheless, calculations based on the foregoing equations will yield results on the high side. This could be attributed to the following reasons:

- 1 As stated earlier, a portion of the deformation energy will not be available as sensible heat and thus contributes nothing to the temperature rise at the shear zone. Taylor and Quinney have reported that about 11 per cent of the energy of deformation remained in a cold-worked mild-steel specimen (21) and 15 per cent in pure (carbonyl) iron (22). Epifanov and Rebinder reported an average value of about 3 per cent in the case of drilling annealed aluminum (23). Recent results on machining a gold-silver alloy at low cutting speed (1 fpm) show that the stored energy amounts

to 5.6 per cent for a shear strain of 1.18 but drops to less than 1 per cent as the strain increases to 3.76 (24). On the other hand, Merchant in his discussion of a paper by Trigger (25) shows that the percentage increase in chip hardness bears a linear relationship to the shearing strain. No saturation is indicated. Since data on the proportion of latent energy in the steel chips produced under speeds and feeds comparable to those used in this paper are virtually nonexistent, the calculations described in the foregoing sections do not take into account that stored energy whose existence cannot be disputed. If a value of 5 per cent were assumed, the calculated mean temperature would be lowered by about 30 deg F.

- 2 Since it is rather difficult to incorporate heat loss at various tool surfaces into the computation, it is completely ignored. While this tends to make the calculated temperature slightly on the high side, the error has been estimated to be small.

- 3 The interface temperature is calculated by adding the temperature rise due to frictional sliding to the mean chip temperature as it leaves the shear zone. This involves an approximation since the temperature along a shear plane is not strictly uniform (26). Analysis shows that under certain conditions the local temperature of the chip in the immediate vicinity of the separating surface could be less than the average.

In view of the uncertainties in data on the thermal properties of work and tool material, the simplifying assumptions made in the analysis, possible experimental inaccuracies and, finally, the moot question of what mean temperature is measured by the tool-work thermocouple, the agreement between the calculated and experimental average temperature is considered good. Even more reliable, is, perhaps, the relative distribution of temperature at the interface as calculated by the present procedure.

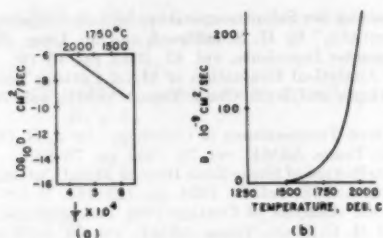
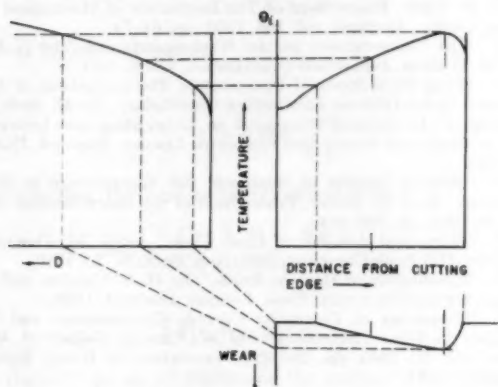


FIG. 14 TEMPERATURE DEPENDENCE OF SURFACE DIFFUSION OF THORIUM IN TUNGSTEN (AFTER LANGMUIR)

(a) SUGGESTED RELATION AMONG TEMPERATURE, DIFFUSION, AND CRATER WEAR



(b) DEVELOPMENT OF CRATER WEAR DURING CUTTING

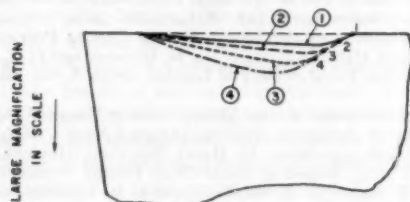


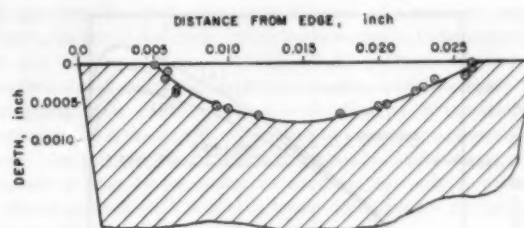
FIG. 15 SCHEMATIC REPRESENTATION OF FORMATION OF CRATER WEAR

TEMPERATURE DISTRIBUTION AT TOOL-CHIP INTERFACE AND FORMATION OF CRATER WEAR

On examining the nature of wear of carbide tools, Dawidl (19, 27) concluded that the cratering type of wear depended mainly on the propensity to "welding" between the chip and the carbide. The latter is connected directly with incipient alloy formation or diffusion. Trent (13) described this as a chemical reaction at the interface.

Experimental data obtained for the interdiffusion of metals exhibit a characteristic feature.¹⁸ When $\log D$ is plotted against $1/T$, straight lines are usually obtained. D is the diffusion coefficient and T the absolute temperature at which the interdiffusion of two dissimilar metals is taking place. The rate of formation of the metallic alloy is solely dependent on D . Fig. 14 illustrates the temperature dependence of surface diffusion of thorium in tungsten. The graph on the right is the same relation when D is

¹⁸ See, for example, "Diffusion," by W. Jost, 1952, Academic Press Inc., New York, N. Y.



Work material: AISI 4150 (mod) annealed, 197 Bhn
Tool material: Steel-cutting-grade carbide
Depth of cut: $u_1 = 0.125$ in.
Feed: $f_1 = 0.00632$ ipr
Cutting speed: $V_c = 456$ fpm
Cutting time: 10.69 min

FIG. 16 CRATER PROFILE DETERMINED BY TOPOGRAPHIC MEASUREMENT

plotted against temperature, on a linear scale. The cratering on the top surface of the tool and its geometrical shape are thought to be due to the nonlinear relationship between D and T and the distribution of temperature at the interface. This viewpoint is illustrated schematically in Fig. 15(a). According to the proposed construction, the lack of symmetry of the crater is evident. This, at first sight, seems to be inconsistent with experience.

Fig. 16 illustrates a crater profile obtained by topographic measurements during the lapping down of the worn tool surface. Flank wear is not shown in the figure. It is seen that the actual profile is more or less symmetrical with respect to a normal to the tool surface. However, closer examination reveals that, as the crater develops, the separation point of the chip and the tool does not stay put but tends to shift along the path 1-2-3-4 as illustrated in Fig. 15(b). The crater profiles during the successive stages of increasing cutting time also are indicated. It is due to this rather peculiar manner of development that the crater profile obtained from a tool after prolonged use appears more or less symmetrical. Further work is in progress to investigate these phenomena.

INFLUENCE OF THERMAL CONDUCTIVITY OF TOOL MATERIAL ON INTERFACE TEMPERATURE

It has been speculated that thermal conductivity of the tool material may play a significant part in influencing the interface temperature and hence on the rate of tool wear. With the method of calculation introduced in this paper, it is a relatively simple matter to examine precisely how variations in K , could affect the temperature of the interface—both the average and the distribution.

Fig. 17 illustrates graphically the results of the calculation. With 50 per cent increase in the value of K , the mean temperature lowers from 1419 to 1395 F (a change of -1.7 per cent). Nevertheless, the proportion of the interface heat which is conducted into tool increases from 9.3 to 13.4 per cent. On the other hand, if the tool conductivity is reduced by half, the mean temperature would increase from 1419 to 1439 F (a change of $+1.4$ per cent). In this case only 4.7 per cent of the interface heat will be transferred into the tool. Obviously, variations in K , have only small influence on the cutting temperature. Of greater significance is the fact that the shape of the temperature-distribution curve is altered to an appreciable extent when K , is varied.

The foregoing statements pertain to the influence of K , on the interface temperature, per se. It is hoped that they may throw some light on the role which conductivity may play in the life of a cutting tool. It is recognized that a change in chemical composition will alter the thermal conductivity of the carbide. However, this effect is secondary to the change in tool-chip adhesion behavior in so far as tool wear is concerned.

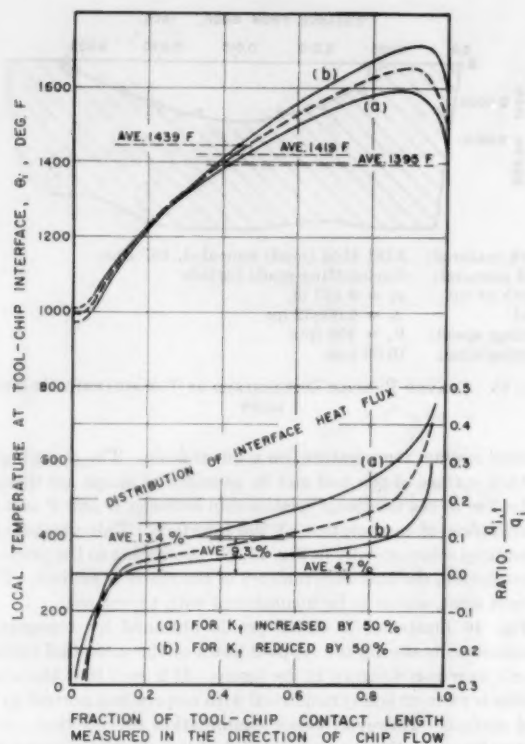


FIG. 17 EFFECT OF THERMAL CONDUCTIVITY OF TOOL MATERIAL ON INTERFACE HEAT FLUX AND TEMPERATURE DISTRIBUTION (CORNER CUTTING)

(Dotted curves are reproduced from Fig. 13.)

ACKNOWLEDGMENT

This paper is based upon an investigation conducted at the University of Illinois and sponsored by the Office of Ordnance Research, U. S. Army, under Contract No. DA-11-022-ORD-1121. The authors hereby express their sincere appreciation to the Office of Ordnance Research for the support of this program. Acknowledgment is made to Kennametal, Inc., Latrobe, Pa., and to Mr. W. L. Kennicott, for the carbide-tool materials used in this investigation; to Mr. Y. H. Lee and Mr. D. L. Mykkanen, Graduate Assistants, and to Prof. D. E. McFeron, all of the Department of Mechanical Engineering, University of Illinois, for their help in computation, preparation of drawings, and helpful discussion during the preparation of the manuscript. Acknowledgment also is made to Mr. B. F. Turkovich, Kearney and Trecker Corporation, Milwaukee, Wis., for his contributions during a portion of the project, and to Miss Irene Cunningham for the typing of the manuscript.

BIBLIOGRAPHY

- 1 "On the Art of Cutting Metals," Part 2, section on "Action and Its Wear in Cutting Metals," by F. W. Taylor, Trans. ASME, vol. 28, 1907, pp. 74-86.
- 2 "Tool and Chip Temperatures in Machine Shop Practice," by H. Shore, M.I.T. thesis, 1924.
- 3 "Die Messung der Schneidentemperatur beim Abdrehen von Flusseisen," by K. Gottwein, *Maschinenbau*, vol. 4, 1925, pp. 1129-1135.
- 4 "The Measurement of Cutting Temperatures," by E. G. Herbert, Proceedings of The Institution of Mechanical Engineers, London, England, vol. 1, 1926, pp. 289-329.
- 5 "Über die Bestimmung des Temperaturfeldes beim Spanablauf," by F. Schwerdt, *Zeitschrift des Vereines deutscher Ingenieure*, vol. 77, 1933, pp. 211-216.

6 "Messung der Schnitttemperatur mittels temperaturanzeigender Farbenstriche," by H. Schallbroch and M. Lang, *Zeitschrift des Vereines deutscher Ingenieure*, vol. 87, 1943, pp. 15-19.

7 "An Analytical Evaluation of Metal-Cutting Temperatures," by K. J. Trigger and B. T. Chao, Trans. ASME, vol. 73, 1951, pp. 57-68.

8 "Surface Temperatures in Grinding," by J. O. Outwater and M. C. Shaw, Trans. ASME, vol. 74, 1952, pp. 73-86.

9 "Distribution of Shear-Zone Heat in Metal Cutting," by W. C. Leone, Trans. ASME, vol. 76, 1954, pp. 121-125.

10 "On the Analysis of Cutting-Tool Temperatures," by E. G. Loewen and M. C. Shaw, Trans. ASME, vol. 76, 1954, pp. 217-231.

11 "Electric Contacts," by R. Holm, Hugo Gebers Förlag, Stockholm, Sweden, 1946, pp. 202-205.

"Mechanical Wear," edited by J. T. Burwell, Jr., ASM, 1950, chapter 14.

12 "The Theory of Plasticity Applied to a Problem of Machining," by E. H. Lee and B. W. Shaffer, *Journal of Applied Mechanics*, Trans. ASME, vol. 73, 1951, pp. 405-413.

13 "Some Factors Affecting Wear on Cemented Carbide Tools," by E. M. Trent, Proceedings of The Institution of Mechanical Engineers, London, England, vol. 166, 1952, pp. 64-74.

14 "Die Temperaturen an der Werkzeugschneide," by E. Bickel and M. Widmer, *Industrielle Organisation*, vol. 8, 1951.

15 "Theoretical Study of Temperature Rise at Surfaces of Actual Contact Under Oiliness Lubricating Conditions," by H. Blok, Proceedings of the General Discussion on Lubrication and Lubricants, The Institution of Mechanical Engineers, London, England, 1938, pp. 222-235.

16 "Moving Sources of Heat and the Temperature at Sliding Contacts," by J. C. Jaeger, Proceedings of the Royal Society, NSW, vol. 76, 1942, pp. 203-224.

17 "Numerical Analysis of Heat Flow," by G. M. Dusenberry, McGraw-Hill Book Company, Inc., New York, N. Y., 1949.

18 "Conduction of Heat in Solids," by H. S. Carslaw and J. C. Jaeger, Oxford University Press, London, England, 1950.

19 "Properties of Cemented Carbide Compositions and Their Relation to Wear Resistance," by W. Dawidl, *Zeitschrift Metallkunde*, vol. 32, 1940, pp. 230-325, translation by Henry Bratcher, Altadena, Calif.

20 "Characteristics of Kennametal," catalog 54, published by Kennametal, Inc., Latrobe, Pa., 1953.

21 "The Latent Energy Remaining in a Metal After Cold Working," by G. I. Taylor and H. Quinney, Proceedings of the Royal Society of London, series A, vol. 143, 1934, pp. 307-326.

22 "The Emission of the Latent Energy Due to Previous Cold Working When a Metal Is Heated," by H. Quinney and G. I. Taylor, Proceedings of the Royal Society of London, series A, vol. 163, 1937, pp. 157-181.

23 "Energy Balance of the Metal Cutting Process," by G. I. Epifanov and P. A. Rebinder, *Doklady Akademii Nauk USSR*, vol. 66, 1949, pp. 653-656, translation by Henry Bratcher, Altadena, Calif.

24 "The Energy Stored in Metal Chips During Orthogonal Cutting," by M. B. Bever, E. R. Marshall, and L. B. Tieknor, *Journal of Applied Physics*, vol. 24, 1953, pp. 1176-1179.

25 "Progress Report No. 2 on Tool-Chip Interface Temperatures," by K. J. Trigger, Trans. ASME, vol. 71, 1949, pp. 163-174.

26 "The Significance of the Thermal Number in Metal Machining," by B. T. Chao and K. J. Trigger, Trans. ASME, vol. 75, 1953, pp. 109-120.

27 "Influence of Diffusion and Alloy Formation on Resistances to Wear of Cemented Carbide Compositions," by W. Dawidl, *Zeitschrift Technische Physik*, vol. 21, 1940, p. 44, translation by Henry Bratcher, Altadena, Calif.

Appendix

From the standpoint of the chip, it is seen that

$$\Delta\theta_i = \frac{(1 - \lambda_c)q_c \csc \phi}{v_c c_p} \quad [15]$$

in which

$$q_c = \frac{F_{\phi s}}{J \omega h \csc \phi}$$

If use is made of Jaeger's solution (16), an approximate expression for $\Delta\theta_i$ can be formulated from the standpoint of the workpiece. Thence

$$\Delta\bar{\theta}_s = \frac{1.064\lambda_s q_s}{K_s} \left(\frac{\kappa_s l_s}{v_s} \right)^{1/2}$$

$$= 0.753 \frac{\lambda_s q_s l_s}{K_s \sqrt{L_s}}, \text{ provided that } L_s > 5 \dots \dots \dots [16]$$

Since

$$\frac{v_s}{v_c} = \frac{\cos \alpha}{\cos(\phi - \alpha)}$$

$$L_s = \frac{v_s l_s}{2\kappa_s} = \frac{v_s l_s \csc \phi}{4\kappa_s} = \frac{v_c l_s \cot \phi + \tan(\phi - \alpha)}{\kappa_s} = \frac{R_f \epsilon}{4}$$

Substituting this latter relation into Equation [16] yields

$$\Delta\bar{\theta}_s = 0.753 \frac{\lambda_s q_s l_s \csc \phi}{K_s \sqrt{R_f \epsilon}} \dots \dots \dots [17]$$

Equating [15] and [17] gives

$$\lambda_s = \frac{1}{1 + 0.753 \sqrt{\frac{R_f}{\epsilon}}}$$

From Equation [15] it can be readily deduced that

$$\Delta\bar{\theta}_s = \frac{(1 - \lambda_s) F_s}{9336 w_1 h C \rho} \frac{\cos \alpha}{\cos(\phi - \alpha)}$$

Discussion

H. BLOK.¹¹ As an alternative to the authors' computational method the writer would suggest the method followed in setting up Equations [34] and [35] of his discussion of an earlier paper.¹²

It has not yet been proved possible to solve the integral equation involved in an exact analytical manner, but numerically it can be solved without any essential difficulty. In fact, it can be solved by the well-known procedure of dividing the integration interval (the length of the tool-chip interface) into a sufficient number of segments and setting up an equal number of linear equations for the temperatures representative of the segments. The alternative method would appear to be more straightforward than the authors' at least in the sense that it does not involve any iteration. In any case it does away with the authors' objection to the less exact method that was published by the writer¹² in 1937, reference (15).

R. S. HAHN.¹⁴ This paper brings us closer to a full understanding of the phenomenon of cratering and the authors have done a very considerable amount of excellent work to obtain their most interesting results.

The discussion that follows is in no way critical of the authors' work but considers a basic assumption made at the outset. Following Blok (authors' reference 15) the authors assume a uniform heat source and a variable partitioning of the heat between tool and chip. In sliding contacts, which Blok and later Jaeger (16) considered, the assumption of a uniform source appears to be a good approximation to reality. At the tool-chip interface

it is questionable as to the reasonableness of a uniform source strength, in fact a definitely nonuniform source is to be expected.

If one wishes to discuss a "one-sided" problem where heat flows into a semi-infinite body along a region of the boundary the assumption of a uniform source multiplied by a partitioning function $q_{1,i}/q_i$ is essentially equivalent to assuming a nonuniform source. On the other hand where a "two-sided" problem is to be discussed it may be necessary to consider not only a variable partitioning function but also a nonuniform source strength in order to obtain realistic results.

The authors, by manipulating the partitioning function (assuming a uniform source strength q_i) have managed to get reasonable "compatibility" between tool-face temperature distribution and chip-surface temperature distribution. However, in so doing, they have obtained unreasonable interface heat-flux distributions, namely, distributions which show a large heat flux into the tool in the vicinity of the point of separation. It would seem reasonable to expect the heat flux to the tool to diminish at this point.

In connection with the distribution of the source strength between chip and tool face it should be remembered that the source strength must be proportional to the product of sliding stress and velocity at the interface. Many photomicrographs of chips show a sharp bending of the crystal elongation lines in the vicinity of the tool face indicating a secondary plastic flow. The velocity of sliding at the chip-tool interface would therefore appear to vary from one end of the interface to the other. At the point $(-l)$, Fig. 3 of the paper, one might expect the usual chip-flow velocity v_f . Again, at the point of separation $(+l)$ the sliding velocity must equal the chip flow velocity v_f . At points near the origin $\xi = 0$ (Fig. 3) the sliding velocity must be less than v_f in order to account for the secondary flow. Thus we might expect to have a variation in sliding velocity along the interface.

Next the shear stress along the interface should be considered. The shear stress would be expected to be largest near the cutting edge since the normal stresses are probably greatest there. As the point of separation is approached the normal stress must fall off and the shear stress along with it.

The product of sliding velocity and shear stress or source strength would thus seem to be larger in the region of the cutting edge and to fall to zero as one approached the separation point.

Perhaps the mathematical problem becomes too difficult but some consideration should be given the variation of source strength in attempting to refine Blok's procedure for tool-chip-interface temperature distributions.

E. G. LOEWEN.¹⁵ The authors should be congratulated for an excellent analysis of the tool-chip-interface temperature distribution. Their approach to the problem is new, interesting, and useful; it fills a distinct gap in our knowledge. In a general way the results are in conformity with what we know about cratering and the recent experimental work of Axer.¹⁶ Axer buried small thermocouples in 0.02-in.-diam holes inside the tool, but could not make measurements close enough to the tool-chip interface to check theoretical results in that zone properly.

Lacking detailed information on the pressure or friction-force distribution in the interface the authors make the reasonable assumption that it is uniform. However, we should recognize that a sudden pressure drop at the end of the chip contact zone is unlikely, so that less heat will be generated near the end than else-

¹¹ Professor of Mechanical Engineering, University of Technology, Delft, Holland.

¹² "On the Solution of Reynolds Equation for Slider-Bearing Lubrication—IV," by F. Osterle, A. Charnes, and E. Saibel, *Trans. ASME*, vol. 75, 1953, p. 1122.

¹³ Authors' bibliography (15).

¹⁴ Research Engineer, The Heald Machine Company, Worcester, Mass. Mem. ASME.

¹⁵ Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

¹⁶ "Temperaturfeld und elektro-chemischer Verschleiss am Drehmeissel," by H. Axer, *Industrie Anzeiger*, vol. 76, no. 11, February 5, 1954, p. 20. Also, "Temperature Field of a Turning Tool and the Reactions in the Zone of Contact," by H. Opitz, *Microtecnic*, vol. 8, 1954, pp. 183-188.

where. As a result the true temperature peak probably lies somewhat closer to the cutting edge than shown in Fig. 12 of the paper.

It is impossible to perform an analysis of this kind without making many assumptions. One tacit assumption of the authors is that there is temperature equality between chip and tool at any given point in the contact zone. Although this is not quite true, there seems no reasonable alternative. It is hardly practical to take into account variations of thermal properties with temperature, which are appreciable in low-alloy steels, because the added complication would be tremendous. However, it does seem more appropriate to use values for K and C at or close to the mean value of θ , (1410 F) rather than θ_c (687 F). The result would be an increase in the calculated temperatures. It would be interesting if the authors could estimate the magnitude of this difference. If one uses the same thermal constants as in this paper, the method of authors' reference (10) gives a mean value for θ , of about 1600 F; if extrapolated values from Figs. 7 and 8 are used instead, this temperature is reduced to about 1450 F. The difference between the 1600 F and 1408 F temperatures is interesting because it corresponds exactly to an apparently slight difference in the average value of $q_{s,i}/q$, ratio of 0.122 and 0.111, respectively.

The authors' complaint about lack of data on the thermal properties of so many important materials in this field is quite justified. However, some recent measurements have been made at M.I.T. and will be published shortly. It was interesting to see confirmed here the conclusions of reference (10) regarding the surprisingly small influence of the thermal conductivity of tool materials on interface temperature.

It would seem quite feasible in the analysis of tool temperatures to use an electric or electrolytic analog model to solve for energy and temperature distribution, thereby eliminating the complex solution of Equation [14]. The obvious advantage of analog methods is that one is no longer restricted to tools with zero rake and clearance angles, in fact even the effect of progressive cratering could be studied with relative ease. It may be questioned, of course, whether such refinements are worth while as long as we know so little about the friction-energy distribution.

In the case of orthogonal cutting, temperature variations perpendicular to the chip-flow direction are undoubtedly small enough to be neglected when the $w_1/2l$ ratio is 4.7, as in the present example. However, in corner cutting this ratio is cut in half and some error will be introduced. Except for the added computation effort there seems no reason why such an effect could not be included.

AUTHORS' CLOSURE

The authors wish to express their appreciation to the several discussers for their interest in the paper and for their helpful comments.

Professor Blok's suggestion of solving numerically the integral equation such as Equation [35] of his discussion of a paper by Osterle, et al.,¹⁷ or Equation [7] of the present paper, modified for variable heat source, is certainly a welcome addition to the refinement in the method of calculating temperature distribution over sliding contacts. If his method is followed, the θ 's in the right-hand side of Equation [14] are then expressed in terms of Q 's and, consequently, the latter can be evaluated by solving simultaneous equations without iteration. Aside from this difference, the general approach to the problem remains unaltered.

Dr. Hahn has posed a very pertinent question regarding the distribution of frictional energy at the tool-chip interface. His question is partially answered in Professor Loewen's discussion.

If one were to assume a nonuniform shear-stress distribution along the path of contact it would be logical to consider the peak stress to occur at or near the cutting edge. However, the distribution from there to the point of separation of chip and tool is still unsettled. It may decrease linearly or exponentially with the possibility of a comparatively small sudden drop at the point of separation. The assumption of uniform interface heat source was made simply because of the lack of quantitative information on the manner of friction energy distribution over the length of contact. Qualitatively, one may expect that the peak temperature would occur at a point more remote from the separating edge than that shown in Figs. 13 and 17 of the paper for reasons as pointed out by Professor Loewen. If, however, the true nature of the source distribution were known, the procedure described in the paper could still be used, with no mathematical difficulty whatsoever. The method is, in fact, quite flexible and has been extended successfully to include the effect of rubbing at the tool flank.¹⁸ This calculation shows that rubbing at the tool flank has significant effect on the distribution of interface heat flux particularly in the vicinity of the cutting edge.

One of the problems encountered in the analytical evaluation of tool-chip interface temperature distribution is the judicious selection of the values for the temperature-dependent properties of the chip material. In this paper they are taken at the chip bulk temperature. Professor Loewen favors the use of the average interface temperature. Probably neither is strictly correct. The use of chip bulk temperature was preferred because of the extremely steep temperature gradient known to exist normal to the rubbing surface under practical cutting conditions. Calculations have been made recently for machining AISI 4142 steel (212 Bhn) at a cutting speed of 700 fpm using, respectively, the chip bulk temperature and the arithmetic average of chip bulk and mean interface temperatures as references. The temperature-distribution curve is 2 to 5 per cent lower for the latter case.¹⁹ At present it is difficult to compare the accuracy of either procedure. Perhaps not until the whole problem is re-examined by using tedious numerical calculations, and taking into account the variation of K_s and C with temperature, can the proper reference temperature be established.

It may be pertinent to mention at this point that if one wishes to incorporate variations of thermal properties with temperature into the computation, Professor Blok's alternate solution, which aims at avoiding iteration, fails.

The use of the electric-analog field plotter in the analysis of tool temperature distribution is undoubtedly feasible in principle. In fact, it has been done by Bickel and Widmer as mentioned in the earlier part of the paper.¹⁹ Bickel's method necessitates the accurate determination of the temperature field on the tool side surface in the vicinity of the cutting zone. This is a difficult task to perform because of the limited space with which one has to work. His published results are not considered as reliable.

With the increasing use of digital computers in both private and public engineering laboratories, the solution of Equation [14] becomes a routine procedure, which can be executed in a very short time—usually in a matter of few minutes. Also, the finite-difference procedure, when modified, can be adapted to tools with rake and clearance angles differing from zero. The treatment of problems with irregular boundaries is well known,²⁰ but computation will be far more tedious since the convenience re-

¹⁷ "Temperature and Heat Flux Distribution at Tool-Chip and Tool-Work Interface in Metal Machining," by B. T. Chao, K. J. Trigger, and Y. H. Lee. M.E. Tech. Note ORD 1121-1, Engineering Experiment Station, University of Illinois, May, 1955.

¹⁸ See Bibliography (14).

²⁰ See, for instance, Bibliography (17).

¹⁷ See footnote 12.

sulting from use of image source due to symmetry no longer exists.

When the ratio $w_1/2l$ becomes smaller, the problem, theoretically, should be treated as three-dimensional. Once again, the basic approach outlined in this paper can be extended without

difficulty to include this complex case. However, as Professor Loewen correctly points out, as long as we know little of the interface friction energy distribution, all such refinements are perhaps futile at the present time.

Cutter Design and Application for Face-Milling Cast Iron and Steel

By O. W. BOSTON¹ AND W. W. GILBERT²

This paper summarizes an extensive research program sponsored by the Office of Production Research and Development during World War II. Two previous papers have presented the results of power and cutter life, respectively. These results, relating to cutter design, are averaged for various grades of cast iron, tool materials, feed, depth, width of bar, number of teeth in the cutter, and the cutting speed. General effects of each variable are obtained from final tool shapes for steel and cast iron which are given for recommended practice.

INTRODUCTION

THIS is the third paper^{3,4} giving results of tests to face mill a variety of metals with cutters of various materials, sizes, and shapes. The work was sponsored by the War Production Board under the auspices of the Office of Production Research and Development, through the Manufacturing Engineering Committee of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS. The work was done on a project with the Engineering Research Institute of the University by the staff and in the laboratories of the Department of Production Engineering at the University of Michigan. This is the third and final paper, and gives results on cutter design and application for both carbide and high-speed-steel face-milling cutters when machining steel and cast iron. The variables, for each material, consisted of axial rake angle, radial rake angle, tooth profile, peripheral relief angle, peripheral cutting-edge angle, nose radius, width of chamfer, number of teeth in the cutter, cutter diameter, blunting of the cutting edges, cutting speed, and cutting fluids. For each of the variables, net power and unit net power at the cutter, together with the cutting speed for a 30-min tool life, are given and recommendations are made for the most appropriate cutter design.

CONDITIONS OF TESTS

In these tests, as in those of the reports previously given,^{3,4} a No. 5 HM Kearney and Trecker Corporation knee-type horizontal milling machine was used. A recording wattmeter was attached to the 20-hp motor, which drives the spindle, so that the gross power, hp_g , the tare power, hp_t , and the net power at the cutter, hp_c , were determined. The gross power minus the tare

was multiplied by the net mechanical efficiency of the machine, 80 per cent, to give the net power at the cutter. This power is also expressed as the unit power at the cutter per cubic inch of metal removed per minute, u hp.

The wear on the flank of the tool was used to indicate tool life. This was measured by a Filar micrometer eyepiece on a microscope or by a toolmaker's microscope. Unless otherwise noted, the cutter was considered failed when the flank wear below the side cutting edge reached 0.030 in. This method of determining tool failure applied to tools of all materials. The material machined consisted of bars, 4 in. wide by 2 in. thick and 18 in. in length, unless noted to the contrary. The relationship between the position of the cutter and the work, together with the nomenclature of the cutter angles, is shown in Fig. 1.

RESULTS OF TESTS

Tool Material. Three types of cutting-tool materials were used, molybdenum high-speed steel (6W, 6Mo, 4Co, 1.5Va), cast non-ferrous metal (Crobalt 138), and sintered carbide (K2S). These were the recommended types for the purpose of milling cast iron. A material of the cast-iron variety, having 20,000 psi tensile strength and a Brinell hardness of 170, was face-milled across the 4-in-wide face, taking a depth of cut of 0.100 in. and feed of 0.010 in. per tooth. The shape of the teeth of each cutter was as shown in Fig. 1 for all three tool materials.

Power values for each of the tool materials are shown plotted in Fig. 2 and values are given in each case for a cutting speed of 200 fpm for high-speed steel tools, 400 fpm for the cast non-ferrous metals, and 800 fpm for the sintered-carbide tools.

For the high-speed tools, operating at 200 fpm, the net power at the cutter varies from 0.1 for the sharp cutters to 0.5 for dull cutters. The hp_c per cubic inch per minute, u hp_c , for the high-speed-steel cutters at this cutting speed of 200 fpm varied from 0.46 for the sharp cutter to 1.4 for the dull cutter. When the cast nonferrous tools were used at 400 fpm, the net power at the cutter varied from 0.22 for sharp cutters to 0.86 for the dull cutters. The unit horsepower at the cutter varied from 0.44 to 1.0 for the sharp and dull cutters, respectively. When carbide cutters were used at 800 fpm, the net power at the cutter varied from 0.45 to 1.0, while the unit power at the 800-fpm speed varied from 0.43 for the sharp cutter to 0.75 for the dull cutter. The net power for the three tool materials increased directly with the cutting speed for the sharp cutters, whereas the unit net power at the cutter, at the three different speeds, remained practically constant at 0.45 for the three tool materials. These unit values are summarized in Table 1.

The value of V_{30} , the cutting speed for 30-min tool life to cause a flank wear of 0.030 in. on each of the tool materials, was obtained from the cutting-speed tool-life line for each tool material cutting under the conditions indicated. The values of V_{30} are shown as small circles connected by a dashed line in Fig. 2, indicating 150 fpm for the high-speed-steel tools, 300 for the cast non-ferrous-metal tools, and 700 for the sintered-carbide tools. This shows a great improvement in the tool life of the sintered-carbide tools over that of the other tool materials.

Axial Rake Angle. In face-milling cutters the axial rake angle, corresponding to the back rake of a single-point turning tool, was

¹ Professor of Mechanical Engineering and Production Engineering; Chairman, Department of Production Engineering, University of Michigan, Ann Arbor, Mich. Life Fellow ASME.

² Consultant, Machinability, General Electric Company, Schenectady, N. Y. Mem. ASME.

³ The first paper, "Power Required by Carbide-Tipped Face-Milling Cutters," by W. W. Gilbert, O. W. Boston, and H. J. Siekmann, Trans. ASME, vol. 75, 1953, pp. 1087-1093.

⁴ The second paper, "Cutter Life for Face-Milling Cast Iron," by W. W. Gilbert, O. W. Boston, and H. J. Siekmann, Trans. ASME, vol. 76, 1954, pp. 607-612.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 2, 1954. Paper No. 54-A-51.

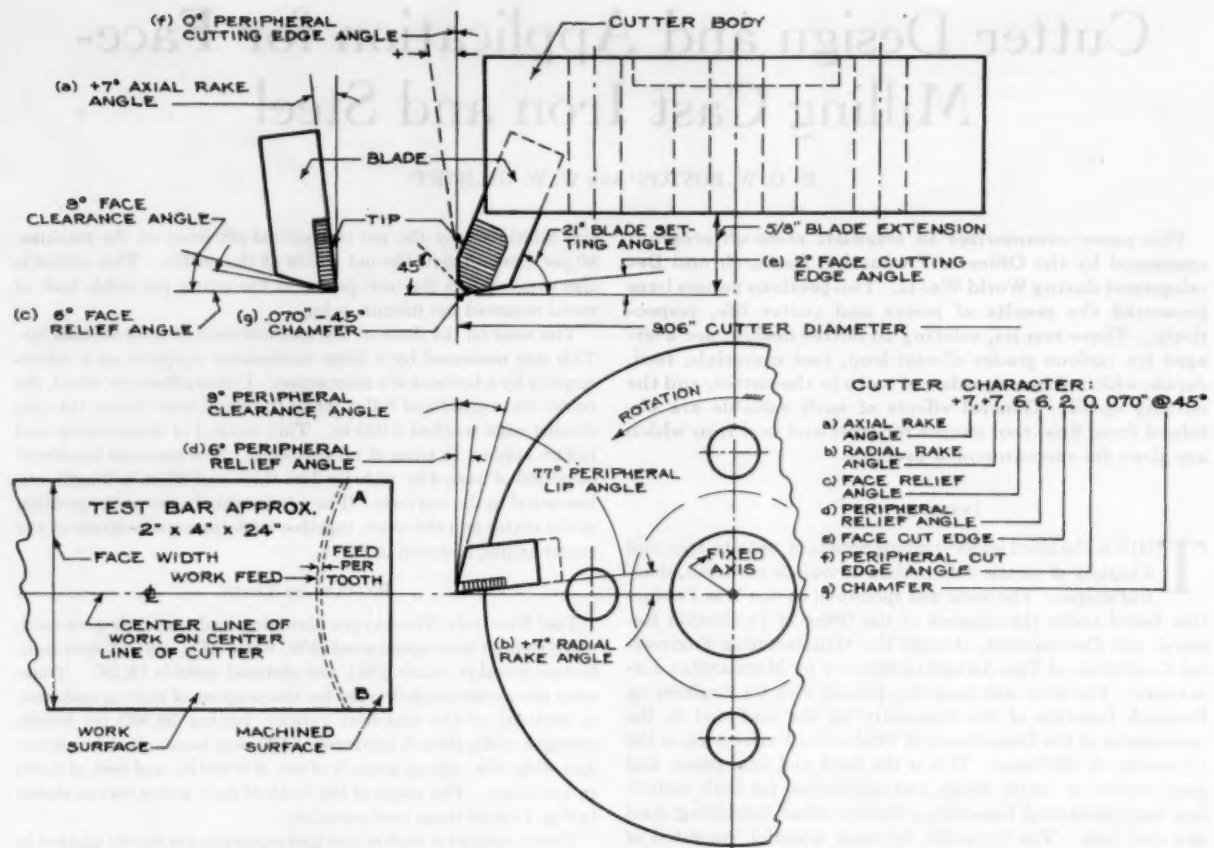


FIG. 1 VIEW OF 9-IN-DIAM FACE-MILLING CUTTER WITH INSERTED BLADE TIPPED WITH CARBIDE. ANGLES OF TOOTH ARE INDICATED WITH A CUTTER CHARACTER SHOWN AT RIGHT. RELATIONSHIP BETWEEN CUTTER AND WORK ALSO IS SHOWN

varied from negative 14 deg to positive 21 deg. For each cutter, the net horsepower at the cutter, hp_c , the unit net horsepower, $u\ hp_c$, and values for V_{30} were obtained when face-milling 40,000-psi (190 Bhn) cast iron, and are shown plotted over the rake angle in Fig. 3. The net horsepower at the cutter, operating at 500 fpm, varied from 0.62 for the negative 14-deg axial rake to 0.50 for the 0-deg rake to 0.48 for the positive 21-deg rake. This gave a reduction in horsepower at the cutter as the rake angle was increased from a negative to a positive value. The unit horsepower for these values of axial rake was 0.62, 0.54, and 0.50, respectively, for the negative 14-deg, 0-deg, and positive 21-deg axial rake angles.

The cutting speed for a 30-min tool life, V_{30} , decreased slightly from 780 fpm for the 14-deg negative rake to 640 fpm for the 21-deg positive rake angle. This shows that negative axial rake angles on the sintered-carbide tool are best for face-milling the cast iron at 190 Brinell hardness.

Radial Rake Angle. Values of horsepower and cutting speed for a 30-min tool life are shown in Fig. 4 for the sintered-carbide face-milling cutter, having various radial rake angles from negative 14 deg to positive 28 deg when face-milling a cast iron of 190 Brinell hardness. The horsepower at the cutter, at a cutting speed of 500 fpm, was 0.57 for the negative 14-deg rake. This value was reduced almost in direct proportion to the change in angle to a value of 0.44 for the positive 28-deg radial rake angle. The unit horsepower also was shown to be reduced in almost direct proportion from 0.6 for the negative 14-deg radial rake angle to 0.49 for the positive 28-deg radial rake angle. The

greatest reduction, however, was from the negative 14-deg to the negative 7-deg radial rake angle.

The cutting speed represented by V_{30} was highest for radial rake angles of 0 to positive 14 deg for which the value was 790 fpm.

When face-milling SAE 4130 steel, having a tensile strength of 110,000 psi and Brinell hardness of 200 to 230 on a 4-in-wide face with a feed of 0.0025 ipt, and a depth of 0.100 in. with K2S carbide-tipped tools, operating at 400 fpm, and the shape used on cast iron, Fig. 3, it was found that the net horsepower at the cutter varied from 0.47 for the negative 14-deg axial rake to 0.38 for the positive 21-deg rake, giving a definite reduction in power. The unit horsepower values are 0.9 and 0.72, respectively. Values of V_{30} increased slightly from 370 fpm for the negative 14-deg rake to 430 fpm for the 21-deg positive rake angle. This shows the positive 21-deg axial rake to be the best for this steel.

Tooth Profile. A study was made of the influence of the tool profile on the power required and the tool life when face-milling a cast iron of 229 Brinell hardness with a K2S sintered-carbide tool of the general shape shown in Fig. 1. The peripheral cutting-edge angle was 45 deg in all cases and the face cutting-edge angle was 2 deg, as shown at A, Fig. 5. The tooth profile was subsequently modified as shown at B and C, respectively, which provided a $1/16$ -in-wide flat at B, and a $1/16$ -in-wide flat with a $1/16$ -in-wide entering angle of $1/4$ deg, shown at C. Cutter A could be considered as a roughing tool, whereas B and C could be considered as finishing tools. The horsepower remained practically constant at 0.2 for all three shapes at the cutting speed of 500

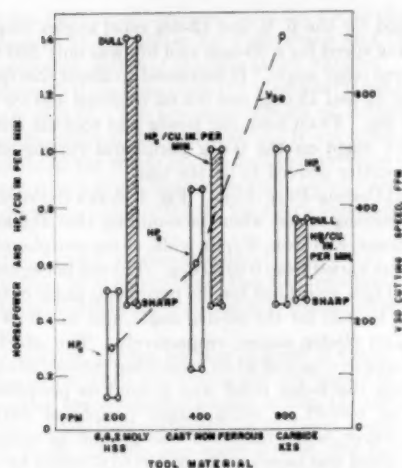


FIG. 2 INFLUENCE OF TOOL MATERIAL ON POWER AND V_{30} WHEN FACE-MILLING CAST IRON HAVING 20,000-PSI TENSILE STRENGTH, 170 BHN, AND 4 IN. WIDTH. DEPTH OF CUT WAS 0.100 IN., FEED 0.010 IPT, AND TOOL SHAPE 7, 7, 6, 6, 2, 0.070 AT 45 DEG

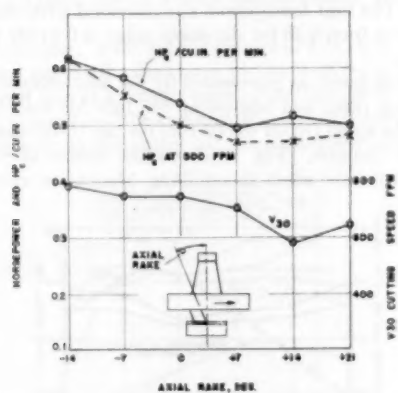


FIG. 3 INFLUENCE OF AXIAL RAKE ANGLE VERSUS HORSEPOWER AND V_{30} WHEN FACE-MILLING 40,000-PSI CAST IRON, 190 BHN. DEPTH OF CUT WAS 0.100 IN., AND FEED 0.010 IPT. WIDTH OF CUT WAS 4 1/4 IN. AND K2S CARBIDE TOOL SHAPE WAS, VAR., 0, 0, 6, 2, 45, 0

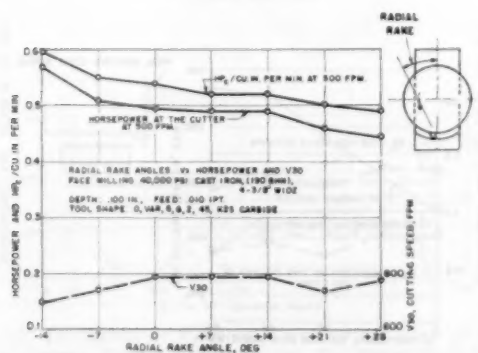


FIG. 4 INFLUENCE OF VARYING RADIAL RAKE ANGLES ON POWER AND V_{30} WHEN FACE-MILLING CAST IRON HAVING A TENSILE STRENGTH OF 40,000 PSI, 190 BHN

fpm. The unit horsepower remained practically constant at 0.50, although that for the cutter A, without the flat, was slightly less at 0.46. The cutting speed represented by V_{30} was about 680 fpm for cutter profiles A and B, but dropped to about 615 fpm for cutter profile C, which had the 1/16-in-wide flat and the 1/16-in-wide entering angle at 1/4 deg. Of the three profiles shown, the cutter A appears to be the best from a tool-life and power point of view.

Peripheral Relief Angle. Variable face and peripheral relief angles were ground on a set of sintered-carbide cutters having a tool shape of 7 deg axial rake and 4 deg radial rake. All cutters

TABLE 1 SUMMARY OF DATA WHEN FACE-MILLING 4-IN-WIDE BARS WITH SINTERED-CARBIDE-TIPPED, SINGLE-TOOTH, 9-IN-DIAM, FACE-MILLING CUTTERS

(The shape of the teeth was +7, +7, 6, 6, 2, 0, 0.070 in. at 45 deg. unless otherwise stated under headings given in the table.)

Variable	Range	Cutting Speed for 30 Min. Life V_{30}	hp_c cu. in. per min. at hp_c (Sharp - Dull)	Cutting Conditions
Tool Material	H.S.S. C.I. Carbide	150 300 700	0.46-1.40 0.46-1.00 0.46-0.75	Cast Iron, 170 BHN $d = 0.100$ in. $f = 0.010$ ipt 0.070 Chamfer
Axial Rake	-11° -7° 0° +7° +11° +21°	780 750 745 710 580 660	0.46 0.58 0.54 0.49 0.52 0.50	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 45° PCRA
Radial Rake	-11° -7° 0° +7° +11° +21°	700 700 790 780 780	0.60 0.55 0.53 0.49	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 45° PCRA
Tooth Profile	2° Face Angle 2° 1/16" Flat 2° 1/16" Flat + 3/4 Lead	680 680 615	0.46 0.50 0.50	Cast Iron, 229 BHN $d = 0.090$ in. $f = 0.010$ ipt 45° PCRA
Relief Angle, Normal	3° 6° 9° 12° 15° 18°	300 310 500 680 680 680	0.60 0.52 0.50 0.47 0.50 0.58	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 0° PCRA
Peripheral Cutting Edge Angle (PCRA) Using 6 Deg. Normal Relief Angle	0° 30° 45° 60°	430 580 660 670	0.55 0.49 0.51 0.48	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt No Chamfer
Rose Radius	0 1/32 1/16 1/8 1/4	430 500 580 660 605	0.45 0.46 0.47 0.51 0.51	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 0° PCRA
Width of Chamfer at 45° PCRA	0 0.010 0.040 0.070 0.120	430 415 415 590 660	0.45 0.49 0.50 0.50 0.51	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 0° PCRA
Number of Teeth	1 2 4 8 16	600 460 450 440 390	0.51 0.50 0.48 0.45 0.45	Cast Iron, 190 BHN $d = 0.100$ in. $f = 0.010$ ipt 0.070 Chamfer at 45°
Cutter Diameter	4 6 8	210 215 220	0.50 0.48 0.45	Cast Iron, 170 BHN $d = 0.100$ in. $f = 0.010$ ipt $w = 1/2$ in. 0.070 Chamfer
Carbide	4 9	360 720	0.44 0.39	
Cutting Fluid	Dry Oil Emulsion	200 305 290	1.0 1.05 1.05	Cast SAE Steel, 4130 $d = 0.100$ in. $f = 0.0025$ ipt
Angle of Blunting	>15 (Sharp) 0° -22-1/2° -45° -67-1/2°	4 81 110 70 23	0.57 0.76 0.67 0.68 0.87	Cast SAE Steel, 4130 $d = 0.100$ in. $f = 0.010$ ipt Dry
Width of Blunting at -22-1/2°	Sharp 0.001 0.005 0.010 0.015	4 21 66 110 105	0.57 0.68 0.67 0.77	Cast SAE Steel, 4130 $d = 0.100$ in. $f = 0.010$ ipt Dry

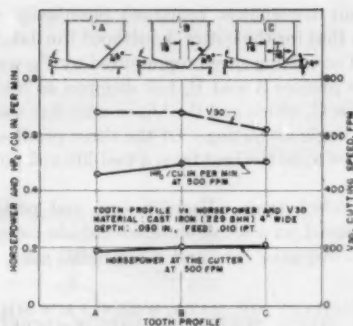


FIG. 5 INFLUENCE OF TOOTH PROFILES ON POWER AND V_{30} WHEN FACE-MILLING CAST IRON, 229 BHN. SINGLE TOOTH WAS OF K2S CARBIDE GROUND TO 7, 7, 6, 6, 2, 45 WITH PROFILES SHOWN

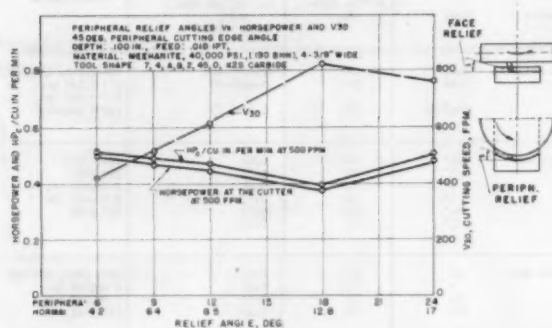


FIG. 6 INFLUENCE OF VARYING PERIPHERAL RELIEF ANGLE WITH CONSTANT PERIPHERAL CUTTING-EDGE ANGLE OF 45 DEG, WHEN FACE-MILLING MEEHANITE HAVING A TENSILE STRENGTH OF 40,000 PSI, 190 BHN

had a 45-deg peripheral cutting-edge angle and a sharp nose at the junction of the end cutting edge and the peripheral cutting edge, as shown at A, Fig. 5. The 40,000-psi Meehanite was face-milled across the $4\frac{1}{8}$ -in. width at a depth of cut of 0.100 in. and a feed of 0.010 in. per tooth. The cutting speed for the power determination was 500 fpm, as shown in Fig. 6. The net horsepower at the cutter and the net unit horsepower were practically the same for this cutting condition, although the lowest value of both was obtained when the peripheral relief was 18 deg measured from the axis and 12.8 deg normal to the cutting edge. The minimum power for this condition was 0.375 net, and 0.4 unit net horsepower. The values of the unit net power for the other cutters were about 0.5.

The cutting speed for a 30-min tool life, V_{30} , is shown to be lowest at about 420 fpm for the 6-deg peripheral relief angle. Its value increases directly with an increase of relief angle to 820 fpm for the 18-deg peripheral relief angle, after which it falls off to 760 fpm for the 24-deg peripheral relief angle. This shows clearly that the maximum tool life and the minimum power occur together for the 18-deg relief angle.

A cutter, using 0-deg peripheral cutting-edge angle instead of the 45 deg referred to previously, was next prepared for test. The peripheral relief angle, which also corresponds to the normal relief angle, in this case, was varied from 3 to 18 deg, as shown in Fig. 7. The net horsepower at the cutter was highest for the 3-deg peripheral relief angle at approximately 0.53. It was about 0.48 for the relief angles of 6 deg, 9 deg, and 12 deg, after which it increased to about 0.51 for the relief angles 15 deg and 18 deg. The hp_c per cubic inch per minute at this speed was 0.6 for the 3-deg peripheral relief angle, but low values of 0.52, 0.5, and 0.47

were obtained for the 6, 9, and 12-deg relief angles, respectively.

The cutting speed for a 30-min tool life was only 200 for the 3-deg peripheral relief angle. It increased to about 620 fpm for relief angles of 12 and 15 deg, and fell off to about 480 for the relief angle of 18 deg. From both the power and tool-life criteria, the 12-deg relief angle on the 0-deg peripheral cutting-edge angle face-milling cutter proved to be the best.

Peripheral Cutting-Edge Angle. Fig. 8 shows the results of net power and cutting speed when face-milling the 40,000-psi (190 Brinell hardness) cast iron, $4\frac{1}{8}$ in. wide, as the peripheral cutting-edge angle was varied from 0 to 60 deg. The net horsepower at the cutter at 300 fpm was 0.315 for the peripheral angle of 0 deg. It was reduced to 0.25 for the 30-deg angle, and was 0.29 and 0.26 for the 45 and 60-deg angles, respectively. The effect of relief angle cannot be overlooked when measuring the tool life or cutting speed. When the 6-deg relief was ground as peripheral relief, there was no benefit in using larger peripheral cutting-edge angles. However, when the relief was ground as normal relief, the cutting speed was increased from 450 to 670 fpm by changing from 0 to 60-deg peripheral cutting-edge angle.

Nose Radius. The standard tool was ground so as to provide various values of nose radii ranging from a sharp point of 0 in. radius to $\frac{1}{4}$ in. radius. The net horsepower at the cutter for the speed of 500 fpm was 0.4 for the sharp nose, 0.445 for the $\frac{1}{32}$ -in. to $\frac{1}{8}$ -in. nose radii, and 0.5 for the largest radius of $\frac{1}{4}$ in., as shown in Fig. 9. The unit horsepower was increased with the increase in nose radius from 0.45 for the sharp point to 0.51 for the $\frac{1}{4}$ -in. nose radius.

The cutting speed, as represented by V_{30} , was lowest at 430 fpm for the sharp point and highest at 640 fpm for the $\frac{1}{8}$ -in. nose radius. The speed fell off to 605 fpm for the $\frac{1}{4}$ -in. nose radius.

Width of Chamfer. Fig. 10 shows the results of power and cutting-speed tests when face-milling Meehanite cast iron of

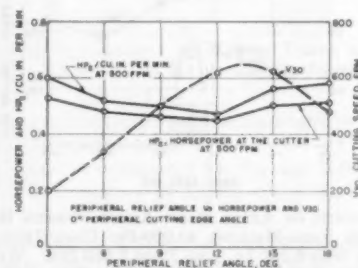


FIG. 7 HORSEPOWER VERSUS PERIPHERAL RELIEF ANGLE FOR CONDITIONS IN FIG. 6, EXCEPT TOOL HAS CONSTANT PERIPHERAL CUTTING-EDGE ANGLE OF 0 DEG

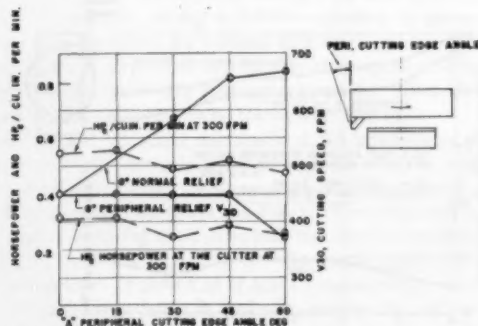


FIG. 8 INFLUENCE OF VARYING PERIPHERAL CUTTING-EDGE ANGLE ON POWER AND V_{30} WHEN FACE-MILLING MEEHANITE HAVING TENSILE STRENGTH OF 40,000 PSI, 190 BHN, WITH K2S TOOTH CUTTER WIDTH OF CUT WAS $4\frac{1}{8}$ IN. AND TOOL SHAPE + 7, + 4, 6, 6, 2, "A".

40,000 psi and 190 Brinell hardness under the standard conditions. The horsepower at the cutter for both sharp and dull tools is shown for different widths of chamfer at the 500 fpm. The sharp-pointed tool with 0 chamfer gives the lowest total net horsepower at the cutter of 0.42. The average value for all cutters, having a chamfer width from 0.010 to 0.150 in., is approximately 0.45. The net power at the cutter for the dull tool averages about 0.66. The unit net power is practically constant for all values of chamfer width, having a value of 0.5 for the sharp tool and 0.65 for the dull tool.

The cutting speed, represented by V_{30} , is low at 420 fpm for chamfer widths of 0.040 in. and less, but increases to 590 fpm for the chamfer width of 0.070 in. and to 660 fpm for chamfer width of 0.150 in.

Number of Teeth. If the number of teeth is varied in the 9-in. face-milling cutter when machining the Meehanite of 190 Brinell hardness under the conditions shown in Fig. 11, the horsepower at the cutter at 500 fpm varies almost in direct proportion to the number of teeth. If only one tooth is used, the horsepower at the cutter is 0.48. It is 3.4 for 8 teeth, and 6.3 for 16 teeth. The unit horsepower at the cutter remains practically constant at 0.45 for all values of teeth, although there is a tendency for a slight increase in unit power for the smaller number of teeth.

The cutting speed for a 30-min tool life, V_{30} , is shown to be highest for the single tooth at 600 fpm, after which it falls off to 480 fpm for 2 teeth, 450 for 4 teeth, 440 for 8 teeth, and 390 for 16 teeth.

Cutter Diameter. Face-milling cutters of 4 in., 6 in., and 9 in., each having one tooth, were used on the 20,000-psi, 170-Brinell hardness cast iron at a cutting speed of 200 fpm, as shown in Fig. 12. Both sintered-carbide and high-speed-steel blades were used. The horsepower at the cutter for the high-speed steel averaged

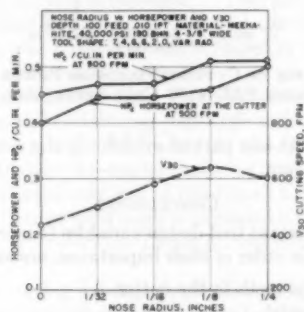


FIG. 9 INFLUENCE OF VARYING NOSE RADIUS ON POWER AND V_{30} WHEN FACE-MILLING MEEHANITE 40,000 PSI, 190 BHN, WITH K2S TOOTH CUTTER

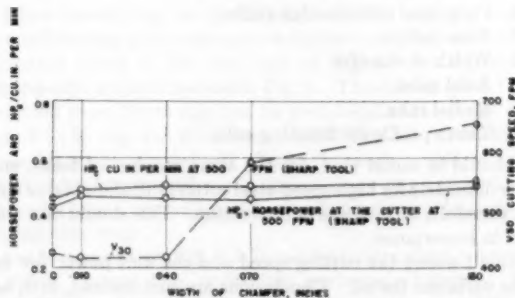


FIG. 10 INFLUENCE OF VARYING WIDTH OF CHAMFER AT 45 DEG FROM CUTTER FACE ON POWER AND V_{30} WHEN FACE-MILLING MEEHANITE CASTINGS, 190 BHN, $4\frac{1}{2}$ IN. WIDE WITH K2S CARBIDE-TIPPED TEETH OF SHAPE 7, 4, 6, 6, 2, 0. DEPTH OF CUT WAS 0.100 IN. AND FEED 0.010 IPT

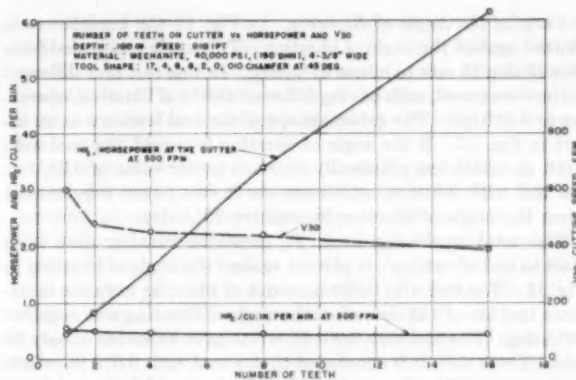


FIG. 11 INFLUENCE OF VARYING NUMBER OF TEETH IN CUTTER ON POWER AND V_{30} WHEN FACE-MILLING MEEHANITE OF 40,000 PSI, 190 BHN, WITH K2S TOOTH CUTTER

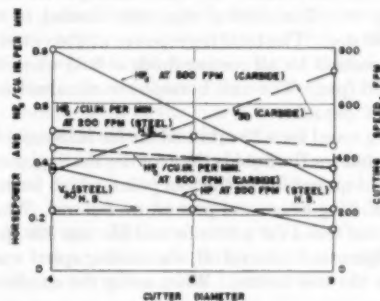


FIG. 12 INFLUENCE OF DIAMETER OF CUTTER ON POWER AND V_{30} WHEN FACE-MILLING CAST IRON HAVING TENSILE STRENGTH OF 20,000 PSI, 170 BHN. DEPTH OF CUT, 0.100 IN., FEED 0.010 IPT AND WIDTH OF CUT WAS 4 IN. TOOL SHAPE 7, 4, 6, 6, 2, 0, 0.070 CHAMFER AT 45 DEG

0.375 for the 4-in.-diam cutter and 0.15 for the 9-in. cutter. The unit power for the high-speed steel cutter, operating at 200 fpm, was 0.5 for the 4-in. cutter, 0.48 for the 6-in.-diam cutter, and 0.45 for the 9-in. cutter. The total power at the carbide cutter operating at 500 fpm was 0.8 for the 4-in. cutter and 0.31 for the 9-in. cutter. The unit power at the carbide cutter operating at 500 fpm was 0.44 for the 4-in. cutter and 0.37 for the 9-in. cutter. All of this shows an appreciable influence in reduction of power as the single-tooth-cutter diameter is increased. However, it also should be noted that the rpm was decreased for the larger cutters.

The cutting speed for a 30-min tool life, V_{30} , for the high-speed-steel cutter remains practically constant at 220 fpm, but increases from 360 fpm to 760 fpm for the K2S carbide-tipped cutter. This shows that both from a power and cutting-speed point of view the high-speed-steel cutters of all diameters are substantially equal. However, for the carbide cutters the larger diameter gives the lowest horsepower, the lowest unit horsepower, and the highest cutting speed, V_{30} .

Fig. 13 is introduced to explain the differences in behavior between the 4-in. and 9-in.-diam cutters when operating at the same cutting speed. The instantaneous maximum values of net power at the cutter for both cutters is the same, but the average value is 0.79 for the 4-in. cutter and 0.31 for the 9-in. cutter. The time of contact, between the tool and the work for the 4-in. cutter, is 0.00105 min and for the 9-in. cutter it is 0.00069 min. This explains very well the difference in their tool life. The 4-in. cutter will take very thin chips at the beginning and end of its cut, whereas the thickness of chip for the large cutter is more uniform throughout the cut.

Varying the Angle of Blunting. In Fig. 14 the horsepower is plotted against the angle of blunting. This angle was varied from plus 16 deg 16 min to minus 67 1/2 deg. In this test four different cutters were used, each having different widths of blunting, operating at 0.010 ipf. The general shape of the tool is shown as an insert in Fig. 15. If the angle of blunting is varied, the tool with 0.001 in. width has practically constant power value of 0.35, but the tool with 0.010 in. width has the lowest power requirement when the angle of blunting is negative 22 1/2 deg.

The total machining time, T_m (machine running time from start to end of cutting), is plotted against the angle of blunting in Fig. 15. The tool with 0.010 in. width of blunting gave the maximum tool life of 145 min when the angle of blunting was negative 22 1/2 deg. The tool with 0.001 in. width gave a tool life of only 20 min. From this, it is concluded that a tool with 0.010 in. width and negative 22 1/2-deg angle of blunting is a good design to give a long tool life with less power.

Cutting Fluids. The influence of cutting fluids, whether dry, oil, or emulsion, was studied when face-milling SAE 4130, as indicated in Fig. 16. The cutting edge was blunted to a width of 0.005 in. at 45 deg. The total horsepower at the cutter remained practically constant for all cutting fluids at 0.35 when the cutting speed was 400 fpm. The unit horsepower remained constant at 1 hp per cu in. per min.

The cutting speed for a 30-min tool life for K2S carbide tool was considerably less for the liquids than for dry cutting, due probably to the frequent quenching of the tool point, which led to the chipping of the cutting edges and premature failure. When cutting dry, the cutting speed for a 30-min tool life was 500 fpm. When using the sulphurized mineral oil, the cutting speed was only 305 fpm to cause the tool failure. When using the emulsion with 20

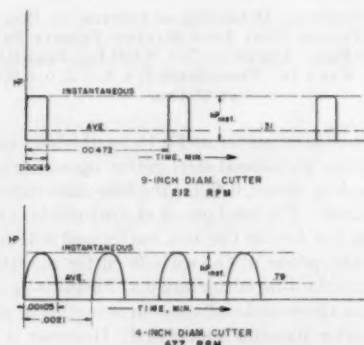


FIG. 13 DIAGRAMS TO SHOW INSTANTANEOUS AND AVERAGE VALUES OF HORSEPOWER FOR 9-IN. AND 4-IN.-DIAM SINGLE-TOOTH FACE-MILLING CUTTERS OF K2S CARBIDE. CUT WAS 4 IN. WIDE, 0.100 IN. DEEP, AND 0.010 IN. FEED PER TOOTH AT A CUTTING SPEED OF 500 FPM

(212 rpm for 9-in.-diam cutter and 477 rpm for the 4-in. cutter. Cutter is on center line of work.)

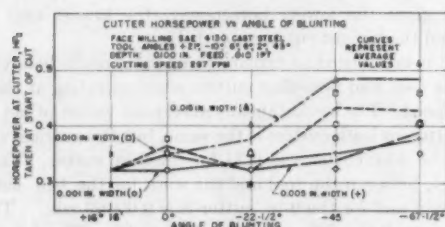


FIG. 14 EFFECT OF CHANGING ANGLE OF BLUNTING ON NET HORSEPOWER REQUIREMENTS WHEN FACE-MILLING 4-IN.-WIDE SAE 4130 CAST STEEL HAVING 200-230 BHN. TOOTH WAS TIPPED WITH K2S CARBIDE

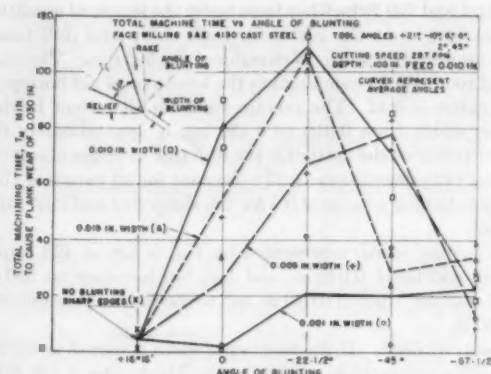


FIG. 15 INFLUENCE OF ANGLE OF BLUNTING FOR EACH OF SEVERAL WIDTHS ON TOOL LIFE WHEN FACE-MILLING SAE 4130 CAST STEEL HAVING 200-230 BHN. TOOTH WAS TIPPED WITH K2S CARBIDE

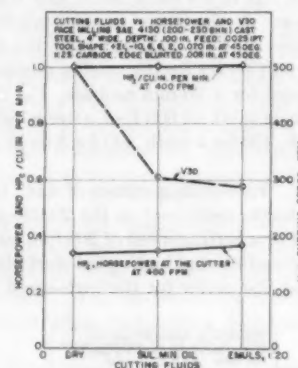


FIG. 16 INFLUENCE OF CUTTING FLUIDS ON POWER AND V_{30} WHEN FACE-MILLING SAE 4130 CAST STEEL, 200-230 BHN

parts of water with one part of soluble oil, the cutting speed was still lower at 290 fpm.

CONCLUSIONS

The most important tool-design variables to increase the rate of metal removal, in order of their importance, are as follows:

- 1 Number of teeth in the cutter.
- 2 Tool material.
- 3 Blunting of the cutting edge—if needed.
- 4 Chipping causing failure.
- 5 Relief angle.
- 6 Peripheral cutting-edge angle.
- 7 Nose radius.
- 8 Width of chamfer.
- 9 Axial rake.
- 10 Radial rake.
- 11 Tooth profile for finishing cuts.

It should be noted that, for the same number of teeth, small cutter diameters for high-speed-steel cutters will allow higher rpm's and, therefore, an increase in production. This should rate about third in importance.

Table 1 shows the cutting speed and the unit power for each of the variables tested. Conclusions are summarized, first, as to net power and unit net power at the cutter and, secondly, as to tool life for the conditions listed for each variable in each figure. The cutting conditions are standardized except as noted.

It is shown that the minimum net unit power, u hp, for the three different tool materials (high-speed steel, cast nonferrous metals,

and sintered-tungsten carbide) remains practically the same at 0.44 for the 20,000-psi cast iron. It was face-milled with sharp cutters for the appropriate cutting speed for each tool material, as shown by the cross-hatched bars in Fig. 2. The unit power for dull cutters may be highest for the high-speed steel inasmuch as this tool material will continue to cut in a duller condition than the cast nonferrous metal or sintered carbide. The latter cutters normally are removed with less wear for reconditioning. The minimum net power at each cutter, hp_c , varies directly with its cutting speed. The value of cutting speed for a 30-min tool life, V_{30} , to cause a flank wear of 0.030 in., for the cutting conditions indicated in Fig. 2, varies as 150, 300, and 700 fpm for the high-speed steel, cast nonferrous metal, and sintered carbide, respectively.

Hp_c and $u\ hp_c$ are reduced as the axial rake angle, Fig. 3, is increased when machining cast iron at 190 Brinell hardness with carbide-tipped face-milling cutters. The lowest values are obtained with the plus 21-deg axial rake angle. V_{30} is decreased as the rake angle is increased from a negative value to a positive 21 deg when face-milling this iron. Therefore the 0-deg to plus 7-deg axial rake seems to be the best for both power and speed.

As the radial rake angle, Fig. 4, is increased from a negative value of 14 deg to a positive value of 28 deg, the hp_c and $u\ hp_c$ at 500 fpm are reduced gradually. Similarly, the cutting speed is increased as the radial rake angle is increased from negative 14 to 0 deg, after which it remains practically constant for all positive values of rake angle up to 28 deg. Therefore a positive radial rake angle of 15 deg is recommended for this 0-deg axial rake.

The profile of a sintered-carbide tool with a 45-deg peripheral cutting-edge angle and a 2-deg face cutting-edge angle has little effect on either power or life. The hp_c and $u\ hp_c$ for the conditions shown in Fig. 5 are increased gradually as the profile is changed by adding a flat to the face of the cutter. They are slightly increased further as a small lead angle of $1/4$ deg is added. V_{30} likewise is reduced as the lead angle is added. Therefore the profile with a sharp nose or small flat is preferred.

As the peripheral relief angle, Fig. 6, is increased on the cutter with a 45-deg peripheral cutting-edge angle and 2-deg face cutting-edge angle, the hp_c and $u\ hp_c$ values are at a minimum when the relief angle is 18 deg (12.8-deg normal relief) when face-milling the 40,000-psi Meehanite cast iron of 190 Brinell hardness. V_{30} is increased as the normal relief angle is increased up to a maximum of 12 to 15 deg. Therefore the 12-deg normal relief angle is recommended.

For a cutter having a 0-deg peripheral cutting-edge angle, Fig. 7, the power is lowest when the peripheral relief angle is from 6 to 12 deg. The cutting speed V_{30} is a maximum for peripheral relief angles of 12 to 15 deg. The 12-deg angle is, therefore, recommended.

When face-milling the 40,000-psi Meehanite cast iron with the face mill having a sintered-carbide-tipped tooth, hp_c and $u\ hp_c$, for a cutting speed of 300 fpm, fall off slightly as the peripheral cutting-edge angle is increased, Fig. 8. The cutting speed V_{30} was increased from 450 to 660 fpm for peripheral cutting-edge angles from 0 to 45 deg, and it increased to 670 fpm for the 60-deg cutting-edge angle. In general, the 45-deg angle would be recommended for commercial work, unless there is an objectionable scale which may require a still further reduction in the peripheral cutting-edge angle.

As the nose radius, Fig. 9, is increased from a sharp point to $1/8$ -in. radius, hp_c and $u\ hp_c$, at 500 fpm, are increased slightly. From a power point of view, smaller nose radii are recommended. The cutting speed, V_{30} , however, is a maximum for the $1/8$ -in. nose radius. It is only slightly lower for the $1/16$ -in. and the $1/4$ -in. radii. From this study a $1/8$ -in. nose radius is recommended.

As the width of chamfer, Fig. 10, on the standard tool is in-

creased from 0 to 0.150 in., the net and unit power is increased only slightly. The cutting speed, V_{30} , however, is increased appreciably for the 0.070 in. width of chamfer and still further for the 0.150 in. The 0.150 in. width of chamfer corresponds to the larger peripheral cutting-edge angle which also proved to be beneficial for tool life.

As the number of teeth in the face-milling cutter, Fig. 11, is increased from 1 to 16, the horsepower at the cutter at 500 fpm increases almost in direct proportion to the number of teeth. The unit net power at the cutter remains practically constant.

Even though V_{30} is highest for the lowest number of teeth, a large number of teeth is recommended to increase the rate of metal removal if sufficient power is available.

As the diameter of a face-milling cutter, having a single tooth and operating at a constant fpm, is increased, Fig. 12, hp_c is reduced appreciably for both carbide and high-speed-steel cutters, and $u\ hp_c$ is reduced slightly. The instantaneous maximum value of power for the larger-diameter cutter equals that of the small-diameter cutter, Fig. 13, but the average power during the cut for the 9-in.-diam cutter is only 0.31 as compared to 0.79 for the 4-in.-diam cutter. V_{30} , however, for the carbide cutter is much higher for the large-diameter cutter than for the small-diameter cutter. V_{30} for the high-speed-steel cutter, however, is practically independent of the cutter diameter. For the carbide teeth, the large-diameter cutter is preferred.

As the angle of blunting of the cutting edge, illustrated in Fig. 14, was varied from plus to minus for different values of width of the ridge, the horsepower became excessive with large negative angles of blunting. For face-milling SAE 4130 cast steel, it is felt that a width of blunting of 0.010 in., at an angle of blunting of negative $22\frac{1}{2}$ deg is best. Fig. 15 shows that the total machining time is greatest for the negative $22\frac{1}{2}$ -deg angle and the 0.010 in. width of blunting.

As cutting fluids were introduced for face-milling the SAE 4130 cast steel, the power remained practically the same for dry cutting, sulfurized mineral oil, and an emulsion, Fig. 16. However, as the liquids were introduced, the cutting speed, V_{30} , dropped from 500 to 300 fpm. Therefore, for face milling steel, with carbide cutters operating at interrupted cuts, dry cutting is recommended.

The best tool angles of each of the individually tested variables cannot always be combined into a given cutter. However, as a result of the tests reported here and other tests, particularly on rake angles for cast steel, two designs of cutters are recommended for sintered-carbide-tipped face-milling cutters, Table 2.

TABLE 2 CHARACTERISTICS OF CUTTER DESIGNS

Cast iron, deg		Steel, deg	
(a) Axial rake.....	0	(a).....	+20
(b) Radial rake.....	+12	(b).....	-10
		(for chip removal)	
(c) Face relief.....	12	(c).....	10
(d) Peripheral relief.....	12	(d).....	10
(e) Face cutting edge.....	2	(e).....	2
(f) Peripheral cutting edge.....	45	(f).....	45
(g) Chamfer 0.100 in. at 45 deg or nose radius of $1/8$ in.		(g) Chamfer 0.100 in. at 45 deg or nose radius of $1/8$ in.	

* For cast iron with scale, a 0-deg peripheral cutting-edge angle is desirable with a large nose radius.

NOTE: For cutting steel, it is advantageous to blunt the peripheral cutting edge to 0.010 in. width and at a negative 22 deg from the cutter axis.

ACKNOWLEDGMENT

This is the third and last paper summarizing the results of the extensive investigation on the milling of several cast irons and steels with cutters of various tool materials. This work was done for the Government in the Department of Production Engineering, University of Michigan, as a project in the Engineering Research Institute. These data have never been published.

Mr. Walter Noffke, Research Engineer, now with the Boeing

Aircraft Company, Seattle, Wash., did much of the original experimental and analytical work. His services are very much appreciated and acknowledged.

Mr. Ananth Mahadevan, a 1954 senior student in the Production Engineering Department, proved very helpful in correlating these data and in the preparation of the drawings.

Discussion

F. W. LUCHT.⁵ The authors are to be complimented for taking the large amount of data they have compiled from the cast-iron milling research program which they conducted during World War II and finally reducing it to a usable form as given in the paper.

The nomenclature of cutter angles in Fig. 1 refers to the chamfer as 0.070×45 deg. In the paragraph referring to the best angles recommended for sintered-carbide-tipped face-milling cutters the chamfer is specified as 0.100 at 45 deg. This raises the question of the correct way for specifying a chamfer. Years of experience indicate the chamfer dimension should always be measured parallel to the chamfer angle because it facilitates the cutter-grinder operator to work up to the given dimension while grinding. When the chamfer measurement is given parallel to the cutter axis, the grinding operation destroys the initial point of measurement. In other words, the chamfer should be specified as 0.070 at 45 deg.

No comment is made about the surface condition of the six sides of the $2 \times 4 \times 24$ -in. cast-iron test bars used for this work. When taking the initial cut, the cutting edge not only contacts scale on the entire cross-sectional area but for the full depth of cut also along the entire 4×24 in. perimeter. On the succeeding cuts the cutting edge is favored because it only contacts scale for the full depth of cut around the 4×24 in. perimeter. The results as obtained from the test runs are not to be questioned because they conform to a certain set of conditions.

Carbide grade K2S is given as "the recommended type for the purpose of milling cast iron." This really confuses the issue because this happens to be a grade of carbide normally recommended for taking the heavier cuts from steel forgings and castings. In actual practice a so-called tungsten-carbide grade is always used for milling cast iron because it resists the abrasive cutting action much better than the steel-cutting grades of carbide which are made with tantalum- and titanium-carbide additive for resisting cratering of the tool face. For example, grade K6 or our competitive grades 883 or 44A are used more frequently for milling cast iron than grade K2S. This in turn will have a definite effect on the conclusions drawn from these tests and may lead falsely to the use of increased rake and relief angles. For example, the recommended 12 deg rake and relief angles have not been favored in the cast-iron milling circles where the more wear-resistant tungsten carbides with slightly lower rake and relief angles have given longer cutter life. Was grade K2S selected to evaluate the performance of a sintered carbide on both cast iron and steel even though it does not conform to the best over-all practice? If this was the case, it would have been interesting to compare the two types of wear resulting from milling each of the two materials.

The selection of the proper face-mill diameter for a given width of cut has a definite effect on the number of pieces per grind obtained from the cutter as the authors have found. We usually recommend a ratio of 8 to 5 for the face-mill diameter to the width of cut or the face-mill diameter is 1.6 times larger than the width of cut. This means for a 4 -in-wide cut we

would use a $6\frac{1}{2}$ to 7 -in-diam face mill or more logically an 8 in. diam because it happens to be a cutter manufacturer's standard and is carried in stock. We would never use a 4 -in-diam face mill to take a 4 -in-wide cut because the cutter tooth enters the cut with a chip having a zero thickness resulting in a highly abrasive cutting action and short cutter life.

Several years ago the writer presented a paper before this Society on the effect of work position when face-milling indicating maximum cutter life was obtained when the tooth entered the cut about 20 deg ahead of the cutter center line in the direction of feed.

A smoother cutting action also was obtained with a setup as used for these tests when the cutter center line was positioned above the work center line, giving a greater arc of cutter contact below the horizontal cutter center line than above it. This would place the direction of the resultant cutting pressure into the fixture and the knee of the machine, and give increased rigidity and longer cutter life.

K. B. KAISER.⁶ The authors are to be complimented on presentation of a most interesting paper on a subject which is inherently very complex because of the large number of variables involved. Since the beginning of attempts at scientific metal cutting, these variables have been a constant headache to the tool designer because, as the paper points out, the best tool angles dictated by one variable do not always correspond to those found best for another variable. Thus, in final cutter design, a compromise is necessary to obtain the best average conditions.

The writer was particularly interested in the results of the tests in which radial and axial rake were varied. While certain trends are indicated by Figs. 3 and 4, it is believed that the authors did not go far enough in reaching the basic angles affecting power and tool life. The radial and axial rake angles are important only in that they may be chosen in combination with the peripheral cutting-edge angle, or bevel, to obtain most desirable effective, or true, rake and shear angles. As is well known, the effective angles at the cutting edge may differ considerably from the apparent axial and radial rakes.

In our experience, the effective cutting angles, as influenced by a correlation of radial and axial rakes and bevel angle, deserve a much higher place on the list of most important tool-design variables than given by the authors.

The data on the effect of varying the peripheral cutting-edge angle suggests a further set of tests to determine whether it is the angle itself or only the influence of the angle on actual chip thickness (feed per tooth remaining constant) which changes the power and tool-life values. It is the opinion of the writer that it is the thinner chip obtained with a 60 -deg angle which causes an increase in the life constant over that for the 0 -deg cutting-edge angle. This appears to be contradicted, however, by the data on power. The power values in Fig. 8 and Table 1 indicate a slight decrease in the power constant with an increase in peripheral cutting-edge angle (and decrease in chip thickness). This is contrary to the writer's experience and other published data by the authors of the present paper which indicate that thinner chips require more power per cubic inch of metal removal.

Fig. 11, showing the effect of number of blades, is interesting. Have the authors an explanation of what might have caused the increase in cutter life for the single-bladed cutter over those with more blades? It appears to the writer that putting more blades in the cutter should merely duplicate the results obtained with one blade. The data show this to be true in the power check. It may be possible that, in the test setup, some other factor such as lack of machine or workpiece rigidity caused vibration which

⁵ Development Engineer, Carboly Department, General Electric Company, Detroit, Mich. Mem. ASME.

⁶ Assistant Chief Engineer, The Ingersoll Milling Machine Company, Rockford, Ill.

reduced the cutter life when using a multiple-bladed cutter.

A comparison between the tool life using a corner bevel and a nose radius probably will come as quite a shock to many people who have long grieved that they could get better life with a radius if only it were not so difficult to grind. The results show little difference between a corner bevel and radius, and what differential there is favors the bevel. We have long maintained that the smooth, unrestricted chip flow produced by a corner bevel, coupled with proper effective rake and shear angles, is much preferable to a radius where chips are crowded toward a focal point and ideal cutting angles are present at only one point on the radius.

The tests on effect of blunting are commendable and should shed considerable light on something that heretofore has been done by guesswork. It has been well known for years among the "old-timers" in metal cutting that a small amount of blunting or "stoning" was beneficial. These are the first data the writer has noted which actually show its effect.

Data of the type given in this paper are helpful in pointing out design factors which should lead to most efficient milling. However, the writer often has a feeling of frustration, as we imagine the authors do also, in the realization that other variables such as scale, sand, hard castings, relative brittleness of carbides, improper cutter grinding, and lack of control of cutter changing,

all play such an important part in final cutter performance. It is seriously questioned that these latter variables, uncontrollable by the cutter designer, might not require considerable further compromise on the angles recommended for cast iron and steel in the paper to obtain best results in everyday high-production milling.

AUTHORS' CLOSURE

The authors very much appreciate the discussions by Mr. Lucht and Mr. Kaiser. Both are in a position to have observed milling practice over a period of years.

Mr. Lucht raises the question as to the use of type K2S carbide in most of this work as being the type not normally recommended. At the beginning of these tests in 1943, a number of grades of carbide of both the steel and cast-iron types were experimented with to determine the type suited for the work. K2S had been recommended, and in our tests it gave wear characteristics both on the flank and by cupping which seemed to be normal and desired. Some grades wore excessively on the flank and on the chamfer and others excessively on the face, but it was felt that the K2S wearing uniformly on both flank and face was a type which could be used satisfactorily in this investigation involving such a wide variety of types of iron and steel to be face-milled.

Dynamics in the Inlet System of a Four-Stroke Single-Cylinder Engine

By C. F. TAYLOR,¹ J. C. LIVENGOD,² AND D. H. TSAI³

This paper presents the results of a recent investigation of the inlet process in a four-stroke, single-cylinder engine with an inlet pipe. The effects of inlet-pipe length, inlet-pipe diameter, inlet-valve timing, and some other design parameters were studied. The results showed that the dynamic effect of the inlet pipe on the engine volumetric efficiency was associated with (a) the process of accelerating the inlet-pipe air column during the first part of the inlet stroke, and taking advantage of the ramming of the air column during the last part of the same inlet stroke; and (b) the action of the standing wave set up in the air column during the preceding cycle. By a suitable choice of inlet pipe and valve timing, a considerable gain in volumetric efficiency over that for zero pipe could be obtained over a wide range of engine speed.

NOMENCLATURE

The following nomenclature is used in the paper:

- A = dimensionless velocity of sound = c/c_0
- A_p = piston area, sq in.
- A_v = effective area of inlet valve, sq in.
- A_3 = cross-section area of inlet pipe, sq in.
- B = bore of cylinder, in.
- C_{av} = average flow coefficient of inlet valve
- c = velocity of sound, ips or fps
- D = diameter of inlet pipe, in.
- e_v = volumetric efficiency (defined by Equation [2] in text)
- k = ratio of specific heat at constant pressure to that at constant volume = 1.4 for air
- L = combined length of inlet pipe and inlet port, in.
- N = rpm
- p = pressure, psia or in. Hg abs (Hga)
- q = ratio of frequency of inlet stroke to fundamental natural frequency of inlet-pipe air column
- r = compression ratio
- S = stroke of piston, in. or ft
- s = piston speed = $2SN$, fpm or fps
- T = temperature, deg F or deg C
- t = time, sec
- U = dimensionless velocity = u/c_0
- u = velocity, ips
- V_c = cylinder volume, cu in.
- V_d = piston displacement volume, cu in.
- X = dimensionless distance = x/L

x = distance from open end of inlet pipe, in.

Z = dimensionless time = ct/L

ρ = density, lbm/in.³

θ = crankangle, deg ATC

Subscripts

- 0 = reference condition in inlet tank outside inlet pipe
- 1 = condition at open end of inlet pipe, immediately after entrance from atmosphere
- 2 = condition at cylinder end of inlet pipe (same as condition at inlet port)
- c = condition in cylinder
- e = exhaust condition

Other notations are defined as they appear in the text.

INTRODUCTION

In the operation of a four-stroke single-cylinder engine, it has been found that a considerable increase in the air flow to the engine, and hence in its power output, can be obtained by the use of a suitable inlet pipe. This simple and effective method of supercharging, often called "inertia supercharging," has been the subject of a considerable amount of investigation in the past 40 years or so, especially from a theoretical point of view (1 to 4).⁴ However, in so far as the authors are aware, no really comprehensive data have been compiled to date, either for the purpose of design or for a basic understanding of the problem.

In the literature, there seem to be two main theories on the process involved in inertia supercharging. One theory is based on the idea of resonance of the inlet pipe (3, 4). It is argued that the amplitude of the "standing wave" at the inlet port would be large if q , the ratio of the frequency of the inlet stroke to the fundamental natural frequency of the inlet pipe, happened to be a simple fraction ($1/2$, $1/3$, etc.). However, in order to realize a gain in air flow, a positive pressure wave must be timed to arrive at the inlet port at the beginning of the inlet stroke. Thus, according to this theory, a gain in air flow would occur when q is equal to $1/2$, $1/3$, and $1/4$. This theory has not been found entirely satisfactory because it does not give adequate account of a number of important experimental phenomena. For example, it does not explain the gain in power output over the case of no inlet pipe when the frequency ratio is different from $1/2$, $1/3$, or $1/4$. In this respect, available data seem to indicate that the difference in power output due to resonance, or lack of it, is often small compared to the difference in power output with or without the use of an inlet pipe.

The other theory is based on the consideration of the inlet pipe as a device for storing kinetic energy (2). It is argued that as air flows toward the cylinder, kinetic energy is stored in the air column by virtue of its velocity. Some of this energy can be recovered toward the end of the inlet process because of the ramming effect of the air column, which tends to build up the pressure at the inlet port and force more air into the cylinder. From this point of view, the importance of the inlet-pipe diameter also becomes readily apparent. This approach is basically sound, but it has not been systematically developed, and the relationship of

¹ Professor of Automotive Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

² Research Engineer, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

³ National Bureau of Standards, Washington, D. C.; formerly, Instructor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

Contributed by the Oil and Gas Power Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 9, 1954. Paper No. 54-A-188.

⁴ Numbers in parentheses refer to the Bibliography at the end of the paper.

the ramming effect to engine design and operating conditions has remained largely unknown. The theoretical solutions worked out in reference (2) on the simplifying assumptions of incompressible and uniform air flow along the length of the inlet pipe have not proved to be in good agreement with experimental data.

The present paper reports the results of a recent investigation of some of the problems of inertia supercharging brought out in the foregoing discussion. In this investigation the effects of the length and diameter of the inlet pipe, as well as the timing and flow capacity of the inlet valve, on the air flow to a four-stroke single-cylinder engine were studied experimentally over a wide range of engine speed. Parallel to the experimental work, a detailed theoretical study based on the assumptions of one-dimensional, compressible, isentropic, and unsteady flow was also made. A relatively large number of solutions was obtained with the aid of a high-speed digital computer. These solutions showed satisfactory agreement with the experimental results.

Owing to the limitation of space, this paper will present primarily the experimental results. Only a brief discussion of the theoretical results will be included here to provide a basis for interpreting the observed phenomena, and for extending the usefulness of the experimental data. Further details of the theoretical investigation may be found in reference (5).

EXPERIMENTAL SETUP AND PROCEDURE

A CFR engine was used in the experimental work. The general test setup is shown schematically in Fig. 1. In the following paragraphs some of the special features of this setup are discussed briefly.

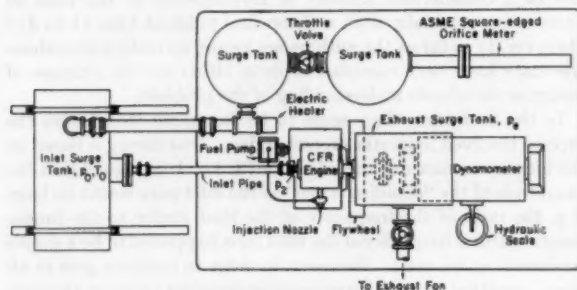


FIG. 1 EXPERIMENTAL SETUP

Inlet System. Air from the atmosphere passed through an ASME orifice meter and a series of three large surge tanks before entering the inlet pipe and the engine. The third tank was movable relative to the engine to permit installation of inlet pipes of different lengths between this tank and the engine. The volume of this tank (dimensions: 14 in. diam, 29 in. length) was approximately 120 times the displacement volume of the engine, and 50 times the volume of the largest inlet pipe (1.50 in. diam, 48 in. length). This size was considered to be large enough so as to have no serious effect on the dynamics of the inlet-pipe air column. Tests at 2400 and 3200 rpm with an inlet pipe of 48 in. length, 1.06 in. diam, all other operating conditions fixed, showed that no measurable change in brake torque occurred when the surge tanks were removed altogether, so that the inlet pipe received air directly from the atmosphere.

Different adapters were used for connecting the third tank to inlet pipes of different diameters. These adapters had well-rounded entrance sections to reduce flow loss. Different adapters also were used for connecting the different inlet pipes to the inlet port. The diameter of the inlet port was $1\frac{1}{8}$ in. For the inlet pipe of 1.50 in. diam, the pipe-to-port adapters had a short ($\frac{1}{2}$

in.) conical section to join the two diameters. For inlet pipes of 1.06 in. and 0.74 in. diam, the adapters had short ($\frac{1}{2}$ in.) flaring sections. In all three cases, since the difference between the diameters of the inlet pipe and the inlet port was not too large, the shape of the transition section was not very important. Prasad and Stahman (6), using the same experimental setup, found that the short transition sections, whether convergent or divergent, had no appreciable effect on the inlet-port pressure.

Valve Timing and the Inlet Valve. Three inlet-valve closing angles, at 33, 60, and 90 deg ABC, were obtained by use of three special cams designated in Table 1 as cams A, B, and C. The timing of the other valve events was fixed. The overlap angle of the inlet and exhaust valves was 6 deg, equally divided before and after top center, and was small enough for the inlet system to be considered isolated from any important dynamic effect of the exhaust system.

Two CFR inlet valves of different designs were tested. Both valves were 1.350 in. OD at the valve head. Their average flow coefficients under steady flow were determined by the method described in reference (7).

Fuel System. The engine operated on the Otto cycle with direct fuel injection into the cylinder beginning at approximately 120 deg ATC on the inlet stroke. The fuel-injection system consisted of a Bosch injection pump with a small (5 mm) plunger, and a wide-angle injection nozzle (Bosch ADN-60-S2 in a Bosch AKB-50-S holder) with a special adapter to fit the cylinder.

Measuring Instruments. The average rate of air flow was measured with a standard 2-in. ASME orifice meter with flange taps across a square-edged orifice of 0.723 in. diam (8). The use of the three large surge tanks in the inlet system was effective in eliminating pressure fluctuations at the orifice so that accurate measurement of the air rate was possible. The average rate of fuel flow was measured by a calibrated flowmeter connected to the injection pump through a suitable surge bellows. Engine speed was measured with a mechanical tachometer and a 60-cycle stroboscope. Brake torque was measured with an M.I.T. hydraulic scale (9).

Pressure in the cylinder was measured with an M.I.T. free-diaphragm balanced-pressure indicator (10, 11). Pressure in the inlet pipe at a point about $1\frac{1}{2}$ in. from the inlet port was measured with a strain-gage-type pressure indicator (12) for the series of tests with cam A, and with the M.I.T. free-diaphragm balanced-pressure indicator for the series with cam C. In each case the diaphragm of the pressure pickup was close to the inside surface of the inlet pipe.

Measurements of the average air rate and fuel rate, engine speed, and brake torque, as well as pressures in the inlet and exhaust tanks and temperature in the inlet tank, involved no particular experimental difficulties. The accuracy of these measurements was considered to be reasonably good. The measured rates of air flow were found to be reproducible to within 2 or 3 per cent.

On the other hand, measurements of pressures in the cylinder and in the inlet pipe were somewhat more difficult. Fig. 2 shows two sets of these pressure records. In some of these records (for example, 1.06-in.-diam pipe at 1600 rpm) the measured inlet-pipe pressure (near the inlet port) was found to be lower than the cylinder pressure during the entire inlet process, although air-rate measurement indicated that air was flowing into the cylinder. This apparent discrepancy may have resulted from instrument error, as well as from the fact that the inlet-pipe pressure was measured at a distance of about 4 in. from the inlet valve. Thus the phase of the pressure wave at the valve was different from that shown in the figure. For example, with the distance of 4 in. between the pressure pickup and the inlet valve, the phase difference could amount to 7 or 8 deg crankangle at 3600 rpm. In spite of these inconsistencies, the measured pressure records were

a few cycles of engine operation until the "steady-state" periodic solution was obtained.

The method just outlined proved to be much too involved for manual computation. Furthermore, the solution to Equation [3] depended on the boundary conditions (configuration of the inlet system and conditions of engine operation), and for each new set of conditions a separate solution would have to be computed. These difficulties made it rather impractical to study the problem extensively by manual computation alone. The method was therefore adapted for computation by a modern high-speed electronic digital computer (15, 16). This made possible the study of a relatively large number of cases, duplicating some of the effects investigated in the experimental phase of this problem.



FIG. 3 TYPICAL SOLUTION BY WWI DIGITAL COMPUTER

Conditions of computation for this solution:
CFR engine at 3600 rpm with cam C, 24-in. inlet pipe, 1.50 in. diam. Other conditions as stated in column 1 of Table 1.

Initial condition was that of quiescent pipe. Computation was carried through one complete engine cycle.

Explanation of Curves:

Horizontal axis represents crankangle. Distance between two neighboring bright spots is 180°. 0° is at extreme left, 720° at extreme right.

Vertical axis is located at crankangle of inlet-valve closing (270° for cam C) and represents dimensionless pressure $(p - p_0)/p_0$ and volumetric efficiency e_v . Distance between two neighboring bright spots is unity for pressure scale and 1/2 for volumetric-efficiency scale, with zero at the horizontal axis.

Top curve represents instantaneous volumetric efficiency. At end of the inlet process this curve shows that volumetric efficiency is 0.83 for this cycle. Middle curve extending entire cycle is pressure wave at inlet port. Bottom curve is cylinder pressure during inlet process.

A typical solution obtained directly from the computer is shown in Fig. 3. The results were obtained in terms of three quantities: inlet-port pressure, cylinder pressure, and "instantaneous volumetric efficiency" defined as

$$\text{Instantaneous volumetric efficiency} = \frac{1}{\rho_0 V_d} \int_{\theta_{IO}}^{\theta} M_a d\theta \dots [1]$$

Here ρ_0 is the reference inlet-air density, V_d the piston displacement volume, M_a the instantaneous mass rate of air flow into the cylinder, and θ the crankangle or time, with θ_{IO} representing the opening crankangle (time) of the inlet valve. Integrated over the entire period of inlet-valve opening, the foregoing expression leads, of course, to the familiar definition of "volumetric efficiency," e_v (17)

$$e_v = \frac{\text{mass of air to cylinder per cycle}}{\rho_0 V_d} \dots [2]$$

Interpretation of Computed Results. Figs. 4 and 5 compare the results of the computed and measured volumetric efficiencies for four series of tests. The numerals alongside the computed points indicate the cycle of computation. For the "first" cycle, the air column was assumed to be initially at rest, with a uniform pressure equal to the inlet-tank pressure. The "second" cycle was

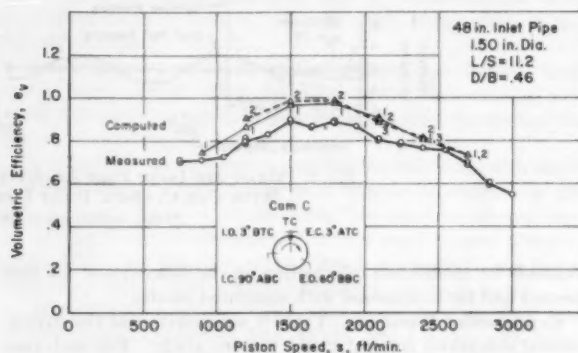
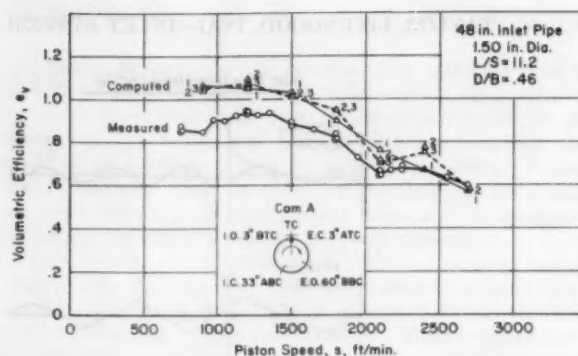


FIG. 4 MEASURED AND COMPUTED VOLUMETRIC EFFICIENCIES FOR TWO VALVE TIMINGS WITH 48-IN. INLET PIPE, 1.50 IN. DIAM (Other conditions as specified in column 1 of Table 1.)

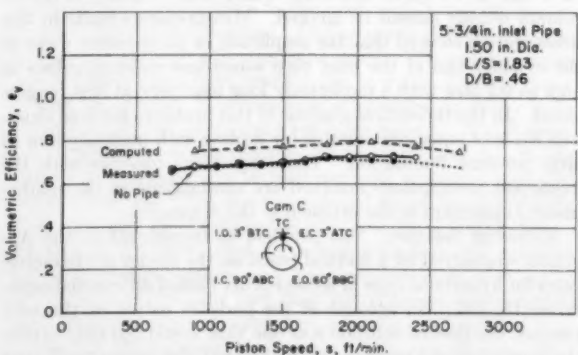
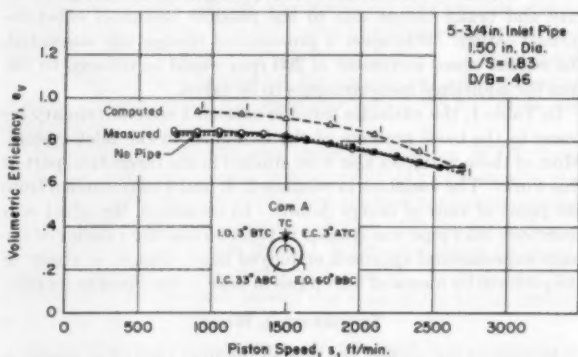


FIG. 5 MEASURED AND COMPUTED VOLUMETRIC EFFICIENCIES FOR TWO VALVE TIMINGS WITH 5 3/4-IN. INLET PIPE, 1.50 IN. DIAM (Measured volumetric efficiencies for cases of zero pipe are also shown. Other conditions as stated in column 1 of Table 1.)

started with pressure waves in the inlet pipe set up as a result of the first cycle. The "third" cycle followed the second cycle in a similar manner. In Fig. 4, the steady-state solution was reached in most cases in the second or third cycle. In Fig. 5, because of the short inlet pipe, the steady-state solution was reached after the first cycle.

In these figures, the measured and the computed (steady-state) values are seen to exhibit similar trends to variation with engine speed, although the computed values are generally higher. This was found to be true in nearly all the cases studied. The difference seemed to be mainly due to the effect of heat transfer (5, 18) which was ignored in the theoretical analysis. In the present setup, the major part of this heat transfer took place not in the inlet pipe but at the inlet valve and in the cylinder. Thus, if the volumetric efficiencies were compared on a "relative" basis, using the volumetric efficiencies for zero (or a very short) pipe as a reference, Fig. 5, the effect of heat transfer would be largely suppressed, while the dynamic effect of the inlet pipe would be automatically brought to the fore. Fig. 6, replotted from the results of Figs. 4 and 5, shows that on this relative basis the dynamic effect in the inlet pipe was predicted with satisfactory accuracy in both series of tests. This also was found to be true in most of the other cases studied.

In order to understand the dynamic effect, it is necessary to have some knowledge of the events taking place in the inlet pipe. In this connection a study of Fig. 7 is of interest. This figure compares the inlet-port and cylinder pressures obtained both by direct measurement and by computation, for the run at 1800 fpm piston speed with cam A, shown in Figs. 4 and 6. Only the first

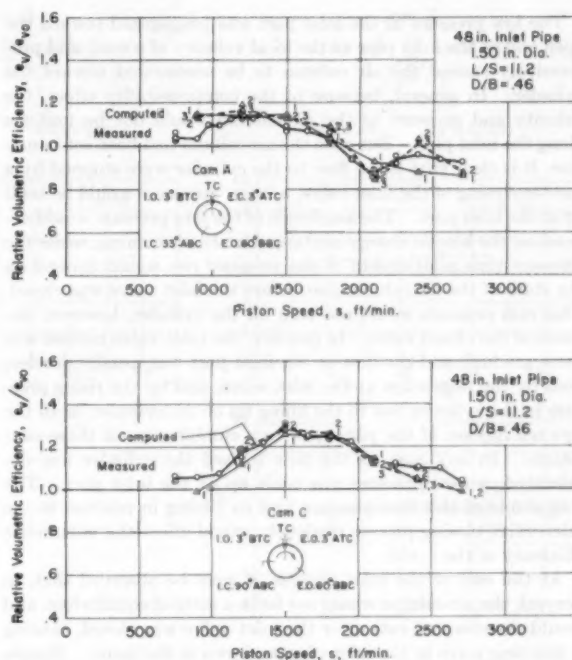


FIG. 6 MEASURED AND COMPUTED RELATIVE VOLUMETRIC EFFICIENCIES FROM FIG. 4 AND 5

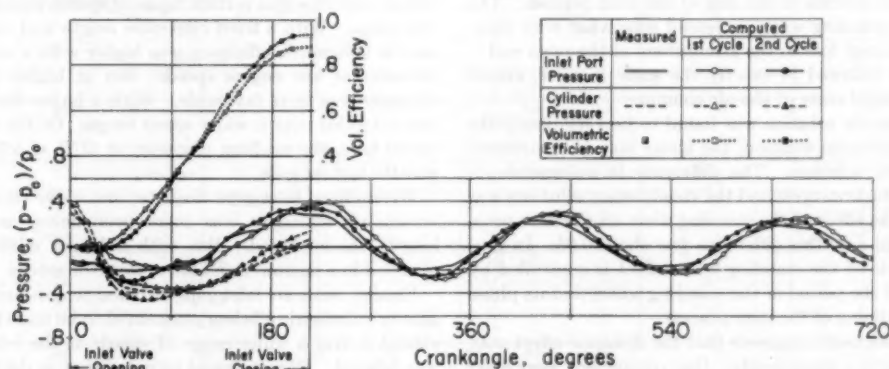


FIG. 7 MEASURED AND COMPUTED PRESSURES AND VOLUMETRIC EFFICIENCIES FOR CASE OF CAM A, 48-IN. INLET PIPE, 1.50 IN. DIAM AT 1800 FPM PISTON SPEED (Other conditions as stated in column 1 of Table 1.)

two cycles of computed pressures are shown here. The pressures of the third cycle are omitted because they turned out to be very nearly the same as those of the second cycle. The volumetric-efficiency curves are the computed instantaneous volumetric efficiencies defined earlier. The measured volumetric efficiency is shown as a single point at the end of the inlet process.

In Fig. 7, the events in the inlet pipe may be described on the basis of the computed results. This should not involve any serious error, since the agreement between the computed and the measured pressures is evidently quite good.

Starting with the inlet stroke of the first computed cycle, it is seen that some outflow from the cylinder took place as the inlet valve began to open. This was because the cylinder pressure was higher than the inlet-port pressure for this particular run. In general, the cylinder pressure at this point is controlled by the

exhaust process. For computation, this pressure was taken from representative experimental data. The outflow caused the inlet-port pressure to rise slightly. This pressure rise was clearly due to the inertia of the air column in the inlet pipe, for if there were no air column, then by assumption (Equation [4] in the Appendix) there would be no pressure rise at this point.

As the inlet process progressed, the cylinder pressure dropped rapidly, due to the combined effect of the outflow and the piston motion away from top center. When the cylinder pressure dropped below the inlet-port pressure, inflow to the cylinder began to take place. At this point the inlet-port pressure began to drop also. This drop was caused by flow velocity as well as the inertia of the air column not yet fully in motion. If there were no air column, the pressure drop would be due to velocity alone, and would be much less.

The low pressure at the inlet port was propagated toward the open end of the inlet pipe at the local velocity of sound, and progressively caused the air column to be accelerated toward the cylinder. In general, because of the compressibility effect, the velocity and pressure of the air column would not be uniform along the inlet pipe. But once the air column had been set in motion, it is clear that if the flow to the cylinder were stopped by a sudden closing of the inlet valve, a "ram" pressure would be built up at the inlet port. The amplitude of the ram pressure would depend on the kinetic energy contained in the air column, while the pressure-time relationship of the pressure rise would depend on the state of the air column just before the inlet valve was closed. This ram pressure would not help fill the cylinder, however, because of the closed valve. In practice, the inlet-valve motion was more gradual, and the flow at the inlet port was usually decelerated by the restriction at the inlet valve, and by the rising pressure in the cylinder due to the filling up of the cylinder, or to the upward motion of the piston, or to a combination of these conditions. In any case, as the flow toward the cylinder was decelerated, a ram pressure was built up at the inlet port. The magnitude of this ram pressure and its timing in relation to the inlet-valve closing process obviously would affect the volumetric efficiency of the cycle.

At the end of the inlet process, it may be observed that, in general, the air column would not be in a state of equilibrium, and would therefore vibrate after the inlet valve was closed, causing a standing wave in the inlet pipe as shown in the figure. Simple calculation would verify that this standing wave was, as expected, at the fundamental natural frequency of the air column. The wave form of the standing wave was apparently determined by the state of the air column at the end of the inlet process. The amplitude of the standing wave decreased somewhat with time, indicating some energy loss to the atmosphere at the open end.

The next cycle followed in exactly the same manner, except with a different initial state of the air column.

Since the third-cycle solution was found to be very nearly the same as the second-cycle solution, the latter may be considered as the steady-state solution. The difference in volumetric efficiency between the first-cycle and the steady-state solutions was evidently due to the effect of the standing wave on the inlet process. In the present case this difference was about 0.05. In general, the magnitude of the standing-wave effect is controlled by the amplitude and the period of the standing wave, and its phase in relation to the timing of the inlet process.

The foregoing discussion suggests that the dynamic effect may be separated into two components: One component associated with the process of accelerating the air column during the first part of the inlet process, and taking advantage of the ramming effect during the last part of the same inlet process; and the other component associated with the action of the standing wave set up in the inlet pipe during the preceding cycle.

In Fig. 6, the component due to ramming is represented by the difference between the first-cycle solutions and the line $e_v/e_{v0} = 1.0$, and the component due to the standing wave by the difference between the steady-state solutions and their corresponding first-cycle solutions. The relative importance of these components depends not only on the dynamic characteristics and the pressure forces on the air column, but also on the design and operating conditions of the engine (e.g., valve capacity and timing, engine speed, etc.). These factors will be further studied by means of the experimental results.

Similarity Rule (Effect of Size). From a mathematical point of view, the condition of dynamic similarity can be established by reference to the governing equations summarized in the Appendix. Since these equations are dimensionless, the solution obtained for one configuration of engine and inlet pipe applies directly to any

of a series of "similar" engines with similar inlet systems under similar operating conditions. From Equations [7] to [7f], it can be seen that for engines and their inlet systems to be similar, L/S , D/B , A_1/A_2 versus θ , and V_c/V_d versus θ must be the same; and for the operating conditions to be similar, p_0/p_c and NL/c_0 must be the same. Since NL is proportional to the piston speed of the engine, it is clear that at the same T_0 (or c_0), similar engines as defined above must operate at the same piston speed (and otherwise the same operating conditions) in order to obtain the same dynamic effect from the inlet pipe and the same relative volumetric efficiency.

These considerations are currently under experimental study at the National Bureau of Standards. Results of this study should provide not only an interesting test of the validity of the similarity rule, but also an indication of departures from this rule due to other factors not accounted for here, such as the effects of heat transfer and fluid friction in the inlet pipe.

DISCUSSION OF EXPERIMENTAL RESULTS

Effects of Inlet-Pipe Length and Diameter and Inlet-Valve Timing. The experimental data on the effects of inlet-pipe length and diameter and inlet-valve timing on the volumetric efficiency of a CFR engine are summarized in Fig. 8. Following the discussion of the similarity rule, these data are presented in terms of the relative volumetric efficiency e_v/e_{v0} against the dimensionless piston speed s/c_0 . The reference curves of e_{v0} versus s/c_0 for the three inlet-valve timings are included at the top of the figure. The following phenomena may be observed:

Generally speaking, in all the cases shown here the use of an inlet pipe resulted in some gain in volumetric efficiency over the case of zero pipe in a certain range of speeds and a loss outside of this range. With a fixed inlet-pipe length and valve timing, the gain in volumetric efficiency was higher with a smaller inlet-pipe diameter at low engine speeds, but at higher speeds a larger diameter was more favorable. With a larger diameter, the gain also extended over a wider speed range. Of the three diameters tested here, the medium diameter of $D/B = 0.33$ was found to give the highest gain.

With a fixed inlet-pipe diameter and valve timing, the gain in volumetric efficiency was more pronounced as the inlet-pipe length was increased. But with a longer inlet pipe, the gain occurred in a narrower and lower range of speeds.

Finally, with an inlet pipe of fixed length and diameter, the gain in volumetric efficiency appeared to be more pronounced and extended over a wider range of speeds as the inlet-valve closing was delayed. Here it should be noted that as the inlet-valve closing angle was changed, the e_{v0} curves were changed also. Thus a high gain in volumetric efficiency (high e_v/e_{v0}) did not necessarily correspond to a high value of actual e_v . Of the three cams tested, the highest e_v was obtained with a moderate inlet-valve closing angle, that of cam B.

In addition to the foregoing, it is seen that for $L/S = 11.2$, $D/B = 0.46$ and 0.33 , under cam A, the volumetric-efficiency curves exhibited some rather pronounced local depression and rise at speeds greater than $s/c_0 = 0.025$. These local variations were evidently due to the effect of the standing wave, as can be seen from a comparison of the corresponding experimental curves in Figs. 8 and 6 with the computed solutions of Fig. 6. Also, with a fixed inlet-pipe diameter and valve timing, the local variations were smaller as the length of the inlet pipe was decreased; and with an inlet pipe of fixed length and diameter, the local variations were smaller as the inlet-valve closing was delayed. It is interesting to observe, also, that in regions of maximum gain in volumetric efficiency, the curves of Fig. 8 showed only small local variations, indicating that, in these regions, the effect of the standing wave was small.

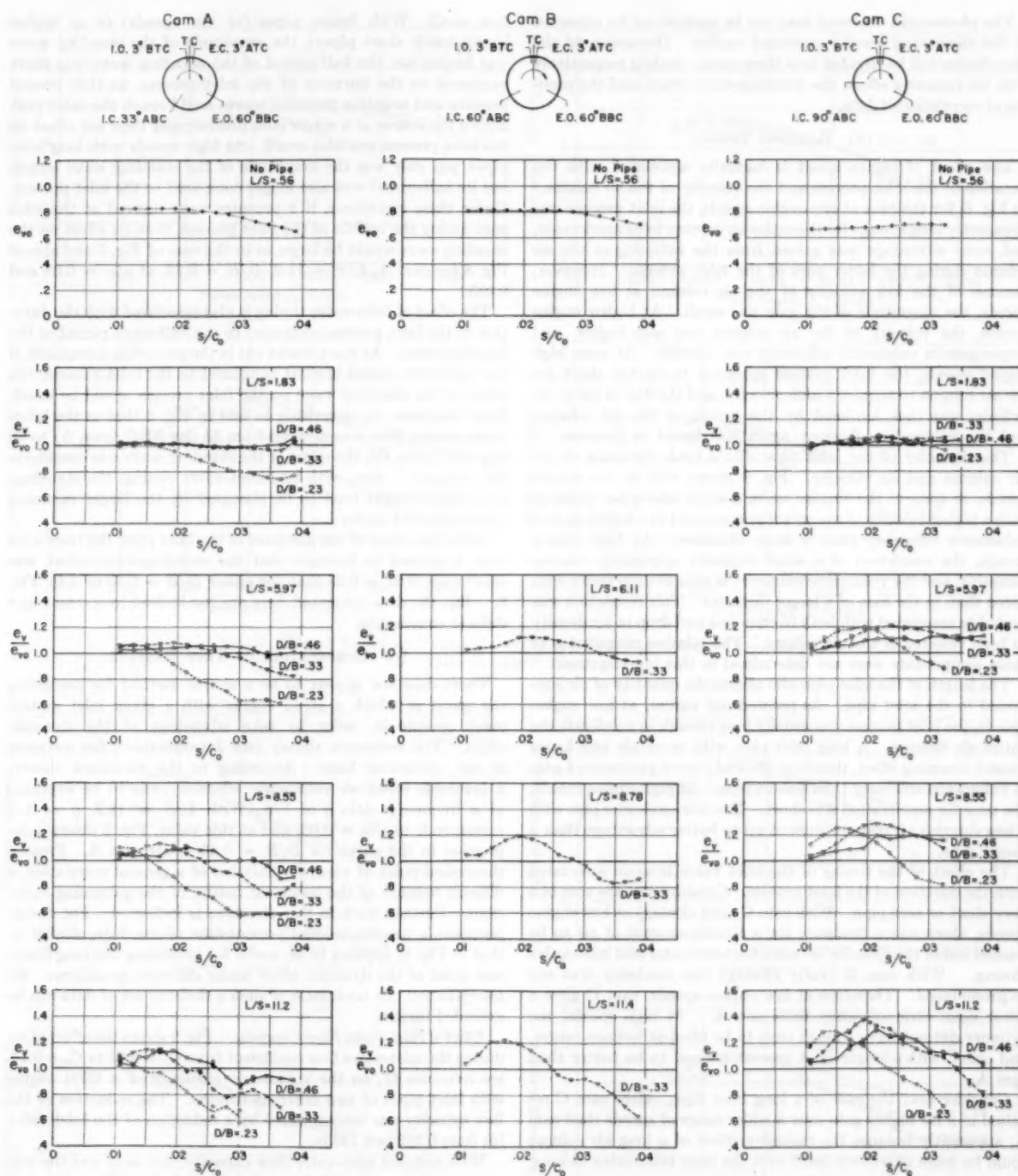


FIG. 8 EFFECTS OF INLET-PIPE LENGTH AND DIAMETER AND INLET-VALVE TIMING ON VOLUMETRIC EFFICIENCY
(Inlet system and operating conditions as stated in column 1 of Table 1.)

The phenomena observed here can be understood by reference to the theoretical results presented earlier. Discussion of the phenomena will be divided into three parts, dealing respectively with the ramming effect, the standing-wave effect, and the problem of correlation of data.

(a) RAMMING EFFECT

The effect of engine speed is basically associated with the duration of the inlet process and the velocity of the air column.⁵ In Fig. 8, for the runs at low engine speeds, the inlet process was apparently long enough for the entire air column to be accelerated, and some advantage was gained from the ramming of the air column during the latter part of the inlet process. However, because of the low velocity of the air column at low engine speeds, the magnitude of the gain was small. At higher engine speeds, the velocity of the air column was also higher, and larger gain in volumetric efficiency was possible. At very high engine speeds, the inlet process appeared to be too short for the air column to acquire a high velocity, and the flow of air to the cylinder was thus hindered by the inertia of the air column, and the volumetric efficiency again was found to decrease.

The diameter of the inlet pipe affects both the mass of the air column and its velocity. Fig. 8 shows that at low engine speeds, in spite of the smaller mass, a small inlet-pipe diameter with a higher velocity of the air column resulted in a higher gain in volumetric efficiency than a large diameter. At high engine speeds, the restriction of a small diameter apparently became excessive, and the volumetric efficiency is seen to drop faster with speed than in the case of a larger diameter. This restriction was probably associated with both friction loss and drop in air density at high velocity of the air column. The relative magnitudes of these components were not determined in this investigation.

The length of the inlet pipe also affects the quantity of air contained in the inlet pipe. As pointed out earlier, at low engine speeds, the inlet process was usually long enough to accelerate the entire air column. A long inlet pipe with more air and hence greater ramming effect, therefore, showed a more pronounced gain in volumetric efficiency than a short pipe. At high engine speeds, the time for acceleration was short. Therefore a shorter pipe with a less sluggish air column showed up to better advantage than a longer pipe.

The effect of the timing of the inlet valve is again associated with the duration of the inlet process. Consider first the case of a very short or zero pipe. With cam C (late closing) at low engine speeds, there was a tendency for a sizable amount of air to be pushed out of the cylinder between bottom center and inlet-valve closing. With cam A (early closing) this tendency was not so pronounced. Therefore at low engine speeds, cam C gave a lower volumetric efficiency than cam A. At high speeds, the cylinder did not have enough time to be filled at bottom center, and cam C with a longer inlet process seemed to be better than cam A.

Consider next the case of a long inlet pipe. Here cam C resulted in a far higher gain over a wider range of speeds than cam A, apparently because the ramming effect of a long air column could be more effectively used with the later inlet-valve closing.

(b) STANDING-WAVE EFFECT

Both engine speed and inlet-pipe length affect the amplitude of the standing wave, but only the pipe length affects its period. In Fig. 8, for the runs at low engine speeds with short inlet pipes, the amplitude of the standing wave was small, and the effect of the standing wave on the inlet process was there-

fore small. With longer pipes (at low speeds) or at higher speeds (with short pipes), the amplitude of the standing wave was larger, but the half-period of the standing wave was short compared to the duration of the inlet process, so that several positive and negative pressure waves could reach the inlet port during the course of a single inlet process, and their net effect on the inlet process was also small. At high speeds with long inlet pipes, not only was the amplitude of the standing wave larger, but its half-period was also longer compared to the inlet process. Under these conditions, if a pressure wave arrived at the inlet port during the middle of the inlet process, then its effect on the standing wave would be large, as in the case of Fig. 7 and also of Fig. 8 for cam A, $L/S = 11.2$, $D/B = 0.46$, at $s/c_0 = 0.03$ and 0.035.

The effect of inlet-valve timing is also associated with the duration of the inlet process compared to the half-wave period of the standing wave. As was pointed out in the preceding paragraph, if the half-wave period is short compared to the inlet process, the effect of the standing wave on the inlet process would be small. It is, therefore, not surprising to find in Fig. 8 that as the inlet-valve closing time was delayed from 33 deg ABC (cam A) to 90 deg ABC (cam C), the effect of the standing wave was considerably reduced. Also, with late inlet-valve closing, the standing-wave effect might tend to be obscured by the larger ramming effect observed earlier.

As for the effect of the diameter of the inlet pipe, the tests with cam A seemed to indicate that the standing-wave effect was smaller for $D/B = 0.23$ than for either $D/B = 0.33$ or 0.46, Fig. 8. But the data presented here are too limited to warrant any definite conclusions.

(c) CORRELATION OF TEST RESULTS

There does not appear to be a simple method for predicting the speed at which a given engine with a given inlet system must operate in order to take advantage of the dynamic effect. The resonance theory (see Introduction), for instance, is not applicable here: According to the resonance theory, a favorable effect on volumetric efficiency was to be expected at a frequency ratio q of 1/3. With $L/S = 11.2$, $q = 1/3$ corresponds to $s/c_0 = 0.03$, and at this value, Fig. 8 shows a depression in the curve for $D/B = 0.46$ under cam A. From a theoretical point of view, formulation of a general correlation is difficult because of the nonlinear nature of the governing equations. Further work in this direction is indicated. For design purposes, a nondimensional presentation of the data, similar to that in Fig. 8, appears to be useful for indicating the magnitude and trend of the dynamic effect under different conditions. By interpolation, the usefulness of such a skeleton set of data can be extended further.

Effect of Inlet-Valve Flow Capacity. Fig. 9 shows the effect of reducing the inlet-valve flow coefficient from $C_{av} = 0.25$ to $C_{av} = 0.14$, see reference 17, on the volumetric efficiency of a CFR engine with inlet pipes of two different lengths. The reduction of the flow capacity was accomplished by a reduction of the inlet-valve lift from 0.262 to 0.145 in.

With reduced inlet-valve flow capacity, not only was the volumetric efficiency reduced, but the amount of gain in volumetric efficiency over the case of zero pipe also was reduced. This was to be expected, since reducing the valve capacity increased the restriction to flow at the inlet valve, and decreased the velocity and the ramming effect of the air column in the inlet pipe.

Effect of Inlet-Valve Design. Fig. 10 compares the volumetric efficiency and the dynamic effect of two inlet pipes due to a change in inlet-valve design. The two valves, A and B, were of quite different configuration, although their average flow coefficients were made the same by adjustment of the valve lift. Under these

⁵ "Velocity of the air column" denotes here the "average" velocity of the air column. In general, the local velocity is not uniform along the length of the inlet pipe.

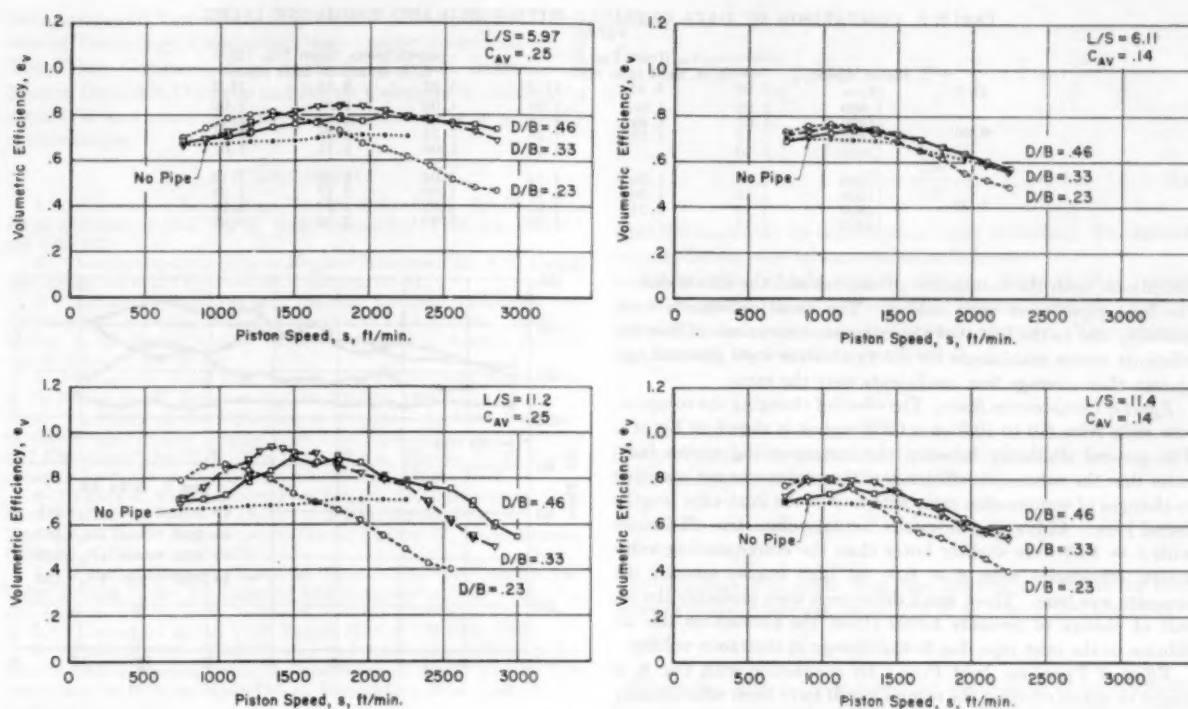


FIG. 9 EFFECT OF INLET-VALVE FLOW CAPACITY ON VOLUMETRIC EFFICIENCY
(Columns 1 and 2, Table 1.)

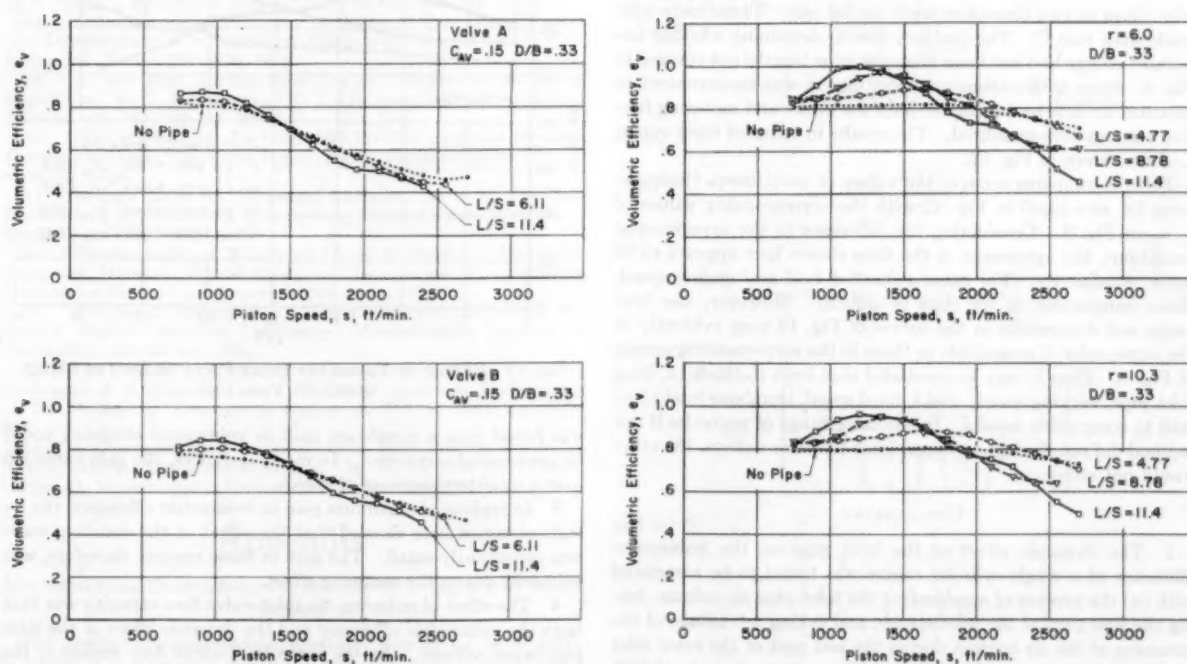


FIG. 10 EFFECT OF INLET-VALVE DESIGN ON CENTER-LINE VOLUMETRIC EFFICIENCY
(Column 3, Table 1.)

FIG. 11 EFFECT OF COMPRESSION RATIO ON VOLUMETRIC EFFICIENCY
(Columns 1 and 4, Table 1.)

TABLE 2 COMPARISON OF DATA OBTAINED WITH FIXED AND TROMBONE INLET PIPES

D/B	Piston speed, fpm	e_p/e_{p0} (from Fig. 8) L/S, fixed inlet pipe			$imep/imep_0$ (from Fig. 12) L/S, trombone inlet pipe		
		5.97	8.55	11.2	5.97	8.55	11.2
0.46	900	1.03	1.06	1.06	1.00	1.01	1.00
	1200	1.07	1.09	1.18	1.02	1.05	1.12
	1800	1.12	1.19	1.25	1.12	1.24	1.26
	2400	1.10	..	1.10	1.06	1.11	1.07
0.33	900	1.08	1.09	1.13	1.05	1.08	1.13
	1200	1.12	1.20	1.29	1.09	1.17	1.23
	1800	1.19	1.24	1.22	1.22	1.24	1.21
	2400	1.13	..	1.00	1.08	1.03	0.92

conditions, both the volumetric efficiencies and the effects due to the inlet pipes were very similar. The small differences were probably due to the fact that the instantaneous values of flow coefficients versus crankangle for the two valves were different, although their average flow coefficients were the same.

Effect of Compression Ratio. The effect of changing the compression ratio from 6.0 to 10.3 on a CFR engine is shown in Fig. 11. The general similarity between the corresponding curves indicates that the volumetric efficiency of the engine was not sensitive to changes of compression ratio with any of the inlet-pipe lengths tested here. At low engine speeds, actual volumetric efficiencies with $r = 10.3$ were slightly lower than the corresponding volumetric efficiencies with $r = 6.0$; at high engine speeds, the opposite was true. These small differences were probably the result of change of pressure forces (from the piston) on the air column in the inlet pipe due to the change in clearance volume.

Effect of Trombone Inlet Pipe. In connection with Fig. 8, it might be asked whether the results would have been substantially different if the inlet pipe, instead of being fixed at a few discrete lengths, was varied in a continuous manner or by small increments.

To settle this question, two series of tests with "trombone" inlet pipes of two diameters were carried out. These tests were made with cam C. The problem was to determine whether important effects had not been missed at pipe lengths not covered in Fig. 8. Since with a trombone inlet pipe it was inconvenient to measure air flow to the engine, only the brake and motoring friction torques were measured. The results in terms of imep versus L/S are shown in Fig. 12.

Table 2 compares some of the values of $imep/imep_0$ ($imep_0 = imep$ for zero pipe) in Fig. 12 with the corresponding values of e_p/e_{p0} in Fig. 8. Considering the difference in the experimental conditions, the agreement of the data shown here appears to be fairly satisfactory. For other values of L/S and engine speed, direct comparison of the data is difficult. However, the local peaks and depressions in the curves of Fig. 12 were evidently of the same order of magnitude as those in the corresponding curves of Fig. 8. Thus it may be concluded that both methods (a, fixed inlet pipe, varying speed; and b, fixed speed, trombone inlet pipe) lead to compatible results. In the same range of variables if one method did not disclose any prominent peaks or valleys, the other would not either.

CONCLUSIONS

1 The dynamic effect of the inlet pipe on the volumetric efficiency of a single-cylinder engine was found to be associated with (a) the process of accelerating the inlet-pipe air column during the first part of the inlet stroke, and taking advantage of the ramming of the air column during the last part of the same inlet stroke, and (b) the action of the standing wave set up in the air column during the preceding cycle.

2 Both the ramming effect and the standing-wave effect were found to depend strongly on the length and diameter of the inlet pipe, the timing of the inlet valve, as well as the speed of the engine. By a suitable choice of the inlet pipe and valve timing, it

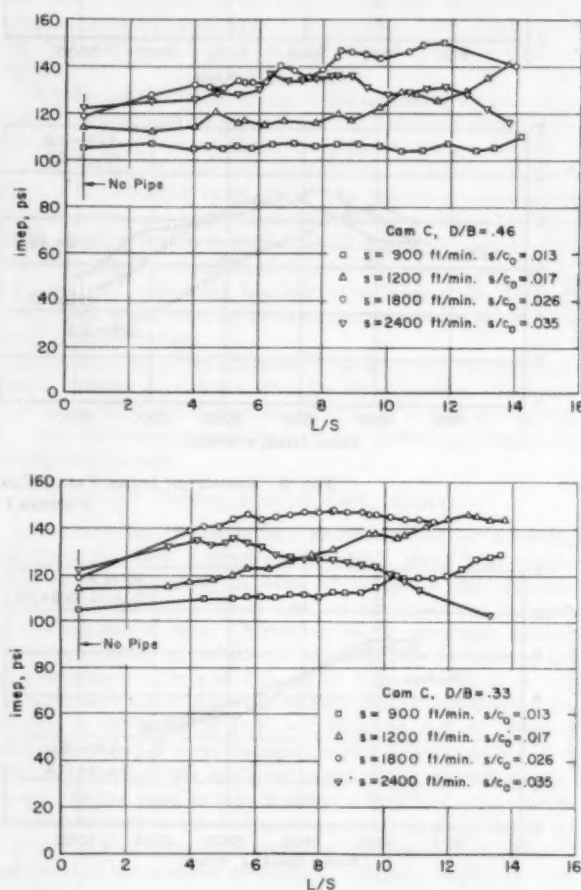


FIG. 12 EFFECT OF TROMBONE INLET PIPES ON ENGINE IMEP (Column 5, Table 1.)

was found that a significant gain in volumetric efficiency could be achieved at any speed. In many instances, the gain extended over a considerable range of speeds.

3 In regions of maximum gain in volumetric efficiency, the results obtained here showed that the effect of the standing wave was consistently small. The gain in these regions, therefore, was primarily due to the ramming effect.

4 The effect of reducing the inlet-valve flow capacity was that both the volumetric efficiency and the dynamic effect of the inlet pipe were reduced. At the same inlet-valve flow capacity, the effects of different inlet-valve design, and of a considerable change in compression ratio, on both the volumetric efficiency and the effect of the inlet pipe, were small.

ACKNOWLEDGMENTS

This investigation was conducted at the Sloan Laboratories

for Aircraft and Automotive Engines at the Massachusetts Institute of Technology, Cambridge, Mass., under the sponsorship of The Texas Company, New York, N. Y. The assistance of Messrs. Donald S. Doremus and Joseph Caloggero in building the apparatus and conducting the experimental tests is gratefully acknowledged.

BIBLIOGRAPHY

- 1 "Resonanzerscheinungen in der Saugleitung von Kompressoren und Gasmotoren," by P. Voissel, *Zeitschrift VDI*, Bd. 56, 1912, pp. 720-722.
- 2 "Inertia Supercharging of Engine Cylinders," by E. S. Dennison, *Trans. ASME, OGP*, vol. 55, 1933, pp. 53-64.
- 3 "Acoustic Vibrations and Internal Combustion Engine Performance. I. Standing Waves in the Intake Pipe System," by P. M. Morse, R. H. Boden, and H. Schecter, *Journal of Applied Physics*, vol. 9, 1938, pp. 16-23.
- 4 "Dynamics in the Inlet System of a Four-Stroke Engine," by R. H. Boden and H. Schecter, NACA TN 935, 1944.
- 5 "A Study of the Dynamics in the Inlet System of a Four-Stroke, Single-Cylinder Engine With Inlet Pipes of Different Lengths and Diameters," by D. H. Tsai, ScD Thesis, Course XVI, Massachusetts Institute of Technology, Cambridge, Mass., 1952.
- 6 "Effect of Varying Inlet Pipe Length and Diameter on Gulp Factor (Inlet Valve Mach Index) of a Four-Stroke Engine," by A. Prasad and R. C. Stahman, SB Thesis, Course II, Massachusetts Institute of Technology, 1952.
- 7 "Effect of Changing Manifold Pressure, Exhaust Pressure, and Valve Timing on the Air Capacity and Output of a Four-Stroke Engine Operated With Inlet Valves of Various Diameters and Lifts," by J. C. Livengood and J. V. D. Eppes, NACA TN 1366, 1947.
- 8 "Fluid Meters—Their Theory and Application, Part I," Report of ASME Special Research Committee on Fluid Meters, THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, New York, N. Y., fourth edition, 1937.
- 9 "Hydraulic Scale for Measurement of Engine Dynamometer Forces," Laboratory Equipment Bulletin No. 6, Diesel Engine Manufacturers Association, January 3, 1949.
- 10 "A New High-Speed Engine Indicator," by E. S. Taylor and C. S. Draper, *Mechanical Engineering*, vol. 55, 1933, pp. 169-171.
- 11 "Improvement of Accuracy of Balanced-Pressure Indicators and Development of an Indicator Calibrating Machine," by J. C. Livengood, NACA TN 1896, 1949.
- 12 "A New High Performance Engine Indicator of the Strain Gage Type," by C. S. Draper and Y. T. Li, *Journal of the Aeronautical Sciences*, vol. 16, 1949, pp. 593-610.
- 13 "The Application of a Graphical Method to Some Dynamic Problems in Gases," by P. De Haller, *Sulzer Technical Review*, No. 1, 1945, pp. 6-24.
- 14 "Lecture Notes on Numerical Analysis," by S. H. Crandall, Department of Mechanical Engineering, Massachusetts Institute of Technology, 1951, chapter 6.
- 15 "Programming for Whirlwind I," by H. Saxenian, Electronic Computer Division, Digital Computer Laboratory, Massachusetts Institute of Technology, Report R-196, June, 1951.
- 16 "Whirlwind I Programmers' Manual (Rough Draft)," by C. W. Adams, Digital Computer Laboratory, Massachusetts Institute of Technology, November, 1951.
- 17 "The Volumetric Efficiency of Four-Stroke Engines," by J. C. Livengood, A. R. Rogowski, and C. F. Taylor, *Quarterly Transactions of The Society of Automotive Engineers*, No. 6, 1952, pp. 617-634.
- 18 "A Thermodynamic Analysis of the Inlet Process of a Four-Stroke Internal Combustion Engine," by C. H. Wu, ScD Thesis, Course II, Massachusetts Institute of Technology, 1947.

Appendix

MATHEMATICAL STATEMENT OF PROBLEM OF DYNAMICS OF INLET SYSTEM

For isentropic, compressible, unsteady flow of air ($p = \rho RT$, $k = 1.4$) in an inlet pipe of uniform diameter, the nondimensional governing equation is a second-order partial differential equation of the hyperbolic type

$$(A^2 - U^2) \frac{\partial^2 \Phi}{\partial X^2} - 2U \frac{\partial^2 \Phi}{\partial X \partial Z} - \frac{\partial^2 \Phi}{\partial Z^2} = 0 \dots [3]$$

where

$$A = \frac{c}{c_0}, \quad U = \frac{u}{c_0}, \quad X = \frac{x}{L}, \quad Z = \frac{c_0 t}{L}$$

and

$$U = \frac{\partial \Phi}{\partial X} \dots [3a]$$

(See Nomenclature for definitions of these symbols.) The boundary conditions may be formulated as follows:

(a) *Open End of Inlet Pipe (Section 1).* When the flow is from the inlet pipe to the atmosphere, the pressure at section 1 may be assumed to be equal to the atmospheric pressure. This assumption leads to the boundary condition

$$A_1 = 1 \dots [4]$$

When the flow is from the atmosphere to the inlet pipe, the flow process may be assumed to be isentropic and "quasi-steady." Then the appropriate boundary condition becomes

$$A_1^2 + \frac{k-1}{2} U_1^2 = 1 \dots [5]$$

(b) *Cylinder End of the Inlet Pipe (Section 2).* When the inlet valve is closed, there can be no flow at section 2, and therefore

$$U_2 = 0 \dots [6]$$

When the inlet valve is open, the assumption of quasi-steady flow may again be used. Considerations of the applicable thermodynamic relationships then lead to the following boundary condition

$$d \left(\frac{p_0 V_s}{p_0 V_d} \right) + (k-1) \left(\frac{p_r}{p_0} \right) d \left(\frac{V_c}{V_d} \right) = k \left(\frac{D^2 L}{B^2 S} \right) \left(\frac{c_2}{c_0} \right)^{\frac{3k-1}{k-1}} (F \text{ or } G) \frac{d\theta}{6NL/c_0} \dots [7]$$

where

$$d\theta = \frac{6NL}{c_0} dZ \dots [7a]$$

and

$$F \text{ or } G = \left[1 + \frac{k-1}{2} M_2^2 \right] M_2 \dots [7b]$$

but for F

$$M_2 = \left[\frac{2}{k-1} \frac{1 - r_2^{\frac{k-1}{k}}}{\left(\frac{A_2}{A_r} \right)^2 \left(\frac{1}{r_2} \right)^{\frac{2}{k}} - 1} \right]^{1/2} \dots [7c]$$

and for G

$$M_2 = - \left[\frac{2}{k-1} \frac{1 - R}{\left(\frac{A_2}{A_r} \right)^2 R^2 - 1} \right]^{1/2} \dots [7d]$$

with

$$R = \frac{1}{2} \left(\frac{1}{r_2} \right)^{\frac{k-1}{k}} \left[1 + \left\{ 1 + 4 \left(\frac{A_r}{A_2} \right)^2 r_2^{\frac{k-1}{k}} \left(r_2^{\frac{k-1}{k}} - 1 \right) \right\}^{1/2} \right] \dots [7e]$$

and

$$r_2 = \frac{p_c}{p_2} \dots \dots \dots [7f]$$

When $r_2 < 1$, F applies; when $r_2 > 1$, G applies; and when $r_2 = 1$, $F = G = 0$.

For the present problem it was desired to find a "periodic" function Φ defined by Equation [3a] which would satisfy Equation [3] and the appropriate boundary conditions as expressed by Equations [4] to [7].

The situation involved here is shown in Fig. 13. In this figure, X represents distance along the inlet pipe, and Z represents time.

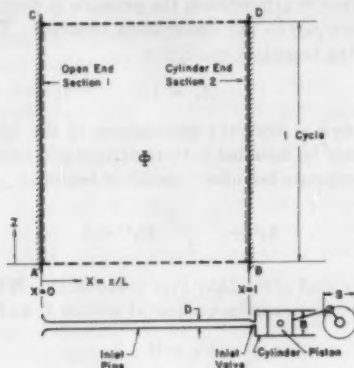


Fig. 13

The desired solution Φ in terms of X and Z must satisfy the boundary conditions along AC and BD , representing, respectively, the open end and the cylinder end of the inlet pipe, while conditions along AB and CD must be identical in the time interval of one cycle of engine operation.

Discussion

T. C. Tsu.⁶ Although the writer has not been working with reciprocating engines for several years, he is glad to accept the invitation to discuss this paper because, some 12 years ago, when he was one of Professor Taylor's students, he played a small part in formulating the basic theory of the inlet and exhaust processes of reciprocating engines.

The Sloan engine laboratory at M.I.T., under the leadership of Professor Taylor, has been persistent in studying the inlet and exhaust processes. The problem investigated 12 years ago was that of a single-cylinder engine with no inlet and exhaust pipes attached whatsoever. A satisfactory theory was developed for predicting the volumetric efficiency and the pressure-time relationships during the inlet and exhaust strokes. The results were subsequently published by the NACA.⁷ The present investigation is an extension of the earlier work, namely, a single-cylinder engine with an inlet pipe attached.

The authors should be congratulated for their success in this endeavor, as evidenced by the good agreement between theory and experiment shown in Figs. 4 to 7, inclusive. This is all the more remarkable when one recalls the assumptions involved in the theory. The governing equation (Equation [3] of the paper) is valid only for one-dimensional, isentropic, potential flow of a

nonviscous gas. In spite of all these restrictions, the theory still offers an adequate explanation of the complicated dynamics of the inlet process. This shows that a little aerodynamic theory goes a long way in rationalizing certain important aspects of engine design. The use of modern digital computers, which are now available to all, transforms the theory into a practical means of exercising design control. In short, the authors have placed at our disposal a method for effecting important savings in time and money when developing new improved engines.

Returning to the governing equation of the theory, the authors are probably aware that their Equation [3] can be made linear with respect to time by making a Legendre transformation. This, however, does not help matters very much because of the complicated boundary conditions. The equations are, perhaps, best solved by means of an automatic computer, such as the authors have done. This should not inconvenience anyone, for the digital computer has become quite popular. Its place in the aviation and automotive industry today is, in the writer's opinion, comparable to that of the desk calculator some 20 or 25 years ago.

In the experimental setup, the authors used a 2¹/₄-cu-ft surge tank before the inlet pipe. This tank-pipe-cylinder combination constitutes a coupled oscillating system. A calculation confirmed the suspicion that the coupling constant was small, even with the shortest and largest pipe. The cylinder and inlet pipe may, therefore, be considered as a simple Helmholtz resonator. Its natural frequency (with any pipe length used by the authors) is fortunately high enough to fall outside the operating range of valve frequency (1/2 of engine rpm). This is consistent with the authors' experimental observation (although made under restricted conditions) that the surge tank had no effect on the inlet-process dynamics. It might be mentioned in passing that with geometrically similar inlet systems operating at constant piston speed, the natural frequency of oscillation is in direct proportion to engine rpm. If model tests reveal no troublesome resonant frequencies, there will be none in the full-size engine.

The authors conclude that the ramming effect is more important than the standing-wave effect in improving the volumetric efficiency. In this conclusion the writer agrees. The practical question is then how to best use this effect in engine design. Among the variables investigated, the writer misses just one, and that is the effect of the rapidity with which the inlet valve is opened and closed. This is a matter of cam design and important limitations exist in practice. Experimentally, therefore, the possible range of investigation is not wide. Theoretically, however, the writer believes it important to determine the limit of possible gains. In other words, what is the maximum benefit that can be derived from the ramming effect? In this respect, the rapidity of valve opening is very important. In the limit, if the valve were able to open and close instantaneously, the inlet pipe would become a shock tube, and one's hopes run high as to the possible gains in volumetric efficiency! Perhaps the authors could tell us what the limit is.

All in all, the authors have presented a very fine paper, a valuable contribution to the understanding of the mechanism of charging a four-stroke single-cylinder engine. The next logical step is to extend the theory to a multicylinder engine feeding from a common intake manifold. Perhaps work along this line is already in progress at M.I.T. or at the National Bureau of Standards. If so, the writer hopes this Society will again have the pleasure of hearing another report from the authors at a future meeting.

AUTHORS' CLOSURE

The authors wish to thank Dr. Tsu for his comments. They agree with Dr. Tsu's appraisal of the modern electronic computer as a tool for engineering research. In this connection, experience

⁶ Formerly, Senior Engineer, Aviation Gas Turbine Division, Westinghouse Electric Corporation, Philadelphia, Pa. Now Research Engineer, Westinghouse Research Laboratories, East Pittsburgh, Pa. Mem. ASME.

⁷ "Theory of the Inlet and Exhaust Processes of Internal Combustion Engines," by T. C. Tsu, NACA TN 1446.

with the present problem showed that the time required for computing the volumetric efficiency and the inlet-port and cylinder pressures for a typical run, Fig. 3, was only about five minutes, while the time required for making experimental measurements of these quantities was approximately 15 minutes. The short computing time was of course due to the fact that WWI was an unusually fast computer. Nevertheless, this comparison illustrates the potentialities of such a computer, not only for making complex theoretical studies, but also for calculating practical design data.

The authors did not investigate the Legendre transformation for the present problem. It was felt that the transformation to what corresponds to the hodograph plane, and the reverse transformation to the physical plane, would be too complicated. In contrast, the method of characteristics is simple in concept and is well-suited for numerical computation on a digital computer because of its stepwise nature.

Regarding the cylinder-inlet-pipe combination as a simple Helmholtz resonator, the authors question the applicability of such an analogy, because of the changing cylinder volume and the inlet-valve area during the inlet process. On the other hand, if the combination of the inlet tank and the inlet pipe is considered as a resonator under the action of the piston in the cylinder, then with an inlet pipe of $5\frac{1}{4}$ in. length and 1.50 in. diam, the

natural frequency of the tank-pipe combination is approximately 17 cps.⁸ This corresponds to 2000 engine rpm (or 1500 fpm piston speed), which is well within the range of speeds covered here. And yet, according to Fig. 5, there was no evidence of resonance either in the computed results or in the measured data. This should not be surprising, since in a Helmholtz resonator the amplitude of the pressure wave is always quite small, even under resonating condition.⁹ Thus the inlet tank indeed served its purpose as a surge tank by keeping the inlet pressure nearly constant.

Dr. Tsu's idea of the limiting case with an inlet valve which opens instantaneously is interesting. In evaluating this scheme, it is important to remember that the average flow capacity of the inlet valve is a controlling parameter.

Thus, to a first approximation, Dr. Tsu's limiting case is roughly equivalent to the conventional case with a higher valve lift, so that its average flow capacity is the same as that of the limiting case. The flow process and the volumetric efficiency cannot be estimated with accuracy. These must be worked out by detailed computation.

⁸ See "Vibration and Sound" by P. M. Morse, McGraw-Hill Book Company, Inc., New York, N. Y., first edition, 1936, pp. 201-202.

⁹ "A Textbook of Sound" by A. B. Wood, G. Bell and Sons Ltd., London, England, second edition, 1949, p. 197.

Steam-Piping Design to Minimize Creep Concentrations

By ERNEST L. ROBINSON,¹ SCHENECTADY, N. Y.

Believing that steam-piping design for high-temperature service has heretofore been based largely upon elastic analysis of expansion stresses without adequate consideration of the importance of high-temperature creep, the author discusses the principles governing the relaxation of expansion stresses during service at high temperature and points out the possibility of creep concentrations in local spots of maximum stress. A number of specific examples are given to show that ordinary piping design usually can be made without such concentrations. Contrariwise, the type of expansion flexibility which invites excessive creep is illustrated. The desirability of cold springing pipe so as to minimize stress at high temperature is emphasized.

A YEAR ago at this time, the Flexibility Section of the ASA Code for Pressure Piping, B31.1, was undergoing revision. The former wording made cold springing to minimize expansion stresses at high temperature permissive, but partially penalized its use by permitting only partial credit for the reduction of stress at high temperature.

A new wording proposed a year ago predicated design upon the range of stress between cold and hot conditions and took the position that self-springing would accomplish as much as cold springing. In so far as the new wording eliminated the partial penalty on cold springing, it appeared to be an improvement, but, on the other hand, the plain implication was that there was no particular advantage in cold springing.

At that time the author of this paper commented on the possibility of self-springing being accompanied by excessive creep in localized regions of high stress. Subsequently the proposed wording was amplified to include a suitable warning about these matters.

The object of this paper is to evaluate the susceptibility to creep of a number of different types of expansion loops and methods of obtaining expansion flexibility and to outline the general principles governing the relaxation of stress due to high-temperature creep in order that proper account may be taken of the plastic strains incurred.

In order to make this subject easily understandable to the ordinary engineer, it seems proper to begin by reviewing the behavior of a simple bolt which drops its stress in proportion to the amount it creeps. In contrast with this simple case, F. P. Coffin's early spring-loaded creep furnace² is discussed where the levers were loaded with long, soft springs instead of weights. He

designed the springs so as to take many times as much creep to accomplish a small amount of stress relief as in the case of a simple bolt. In between these two extremes come the ordinary cases of a bolted flange and the expansion loops of the piping system.

CREEP RELATIONSHIPS

The physical laws governing the plastic deformation of metals under stress for long periods of time at high temperature have been discussed at great length by many writers. The author of this paper proposes to use the simplest relationships which he believes to have generality of application. For this purpose the straight-line log-log plot is well known and widely useful. This is represented by

$$\frac{\tau}{\tau_0} = \left(\frac{S}{S_0} \right)^n \quad [1]$$

The slope of the line is n where $n = \left(\log \frac{\tau}{\tau_0} \right) / \left(\log \frac{S}{S_0} \right)$ and it

differs for different materials. It is likely to be smaller at higher temperatures.

For analysis of behavior where the total amount of creep is limited, it is desirable to use data from tests run with no greater amounts of extension than what is anticipated as likely to occur in use. Relaxation-type tests meet this qualification. Such tests show that there is likely to be a change in slope after a week to a month, after which it remains fairly constant. In other words, after rounding a slight knee the line straightens out for steady, long-time behavior. For refined computations it would be possible to recognize this change in the creep index in the early days of relaxation, but for comparative relationships it does not seem necessary and the value representing the trend taken by the line after the first thousand hours should usually be satisfactory. Relaxation tests interrupted by furnace cooling resume on return to temperature without a break in the line.

It must be recognized that different materials behave quite differently from one another and that suitable values of S_0 and n must be used. It is also true that individual pieces of the same material differ very consequentially among themselves. For the sake of reasonably valid comparisons, the author proposes to use the same relationships in all cases and the same material for making comparisons between similar designs.

Actual piping systems have to meet the complex requirements of a modern power station. It is not the purpose of this paper to analyze any particular piping system. The object is to compare different means for obtaining flexibility. In order to accomplish this, comparisons are made between different series of symmetrical loops whether in bending or in torsion or both.

Only the simplest elements of the problem are included in order to emphasize the importance of carrying on plastic analysis at the same time that elastic analysis is made. Certainly the improvement and extension of the formulations here offered for the computation of plastic deformation by anyone who wishes to undertake it will be most welcome.

¹ General Electric Company. Fellow ASME.

² "The Effect of Temperature on Materials Required in Turbine Design," by S. H. Weaver, General Electric Review, November, 1930, pp. 654-657, Fig. 2.

Contributed by the Power Division and presented at the Annual Meeting, New York, N. Y., November 28-December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 28, 1954. Paper No. 54-A-186.

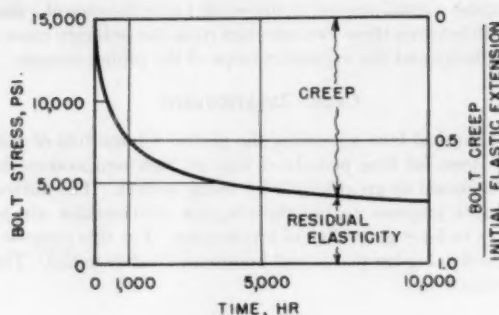
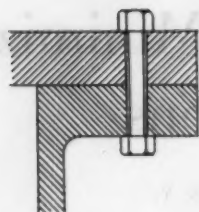


FIG. 1 BEHAVIOR OF BOLT IN AN UNYIELDING FLANGE

(2.25 per cent Cr-1 per cent Mo piping material at 1000 F

$S_0 = 7500$, $E = 20,000,000$, $n = 4$

Initial stress, 15,000; final stress, 3730; total creep, 0.056 per cent; 10,000-hr creep is 0.75 times initial elastic extension.)

BOLT IN AN UNYIELDING FLANGE

This well-known example, Fig. 1, is reviewed in detail here because the relationships form the basis of all the more elaborate applications which follow. The relaxation creep test was made the basis of two progress reports by Project 16 of the ASME-ASTM Joint Committee on the Effect of Temperature on the Properties of Metals.^{3,4}

The initial take-up on the nuts is unchanged thereafter and remains constant so that a formula may be written to show that the elastic extension at any instant plus the previous plastic extension is always equal to the initial elastic extension.

$$\frac{S_1}{E} L = \frac{S}{E} L + pL \text{ or } S_1 - S = Ep \dots \dots [2]$$

Differentiating with respect to time gives the elastic requirement that

$$-\frac{dS}{dt} = E \frac{dp}{dt}$$

But $dp/dt = r$ and the creep properties give

$$r = \frac{r_0}{S_0^n} S^n$$

From which

$$dt = \frac{S_0^n}{r_0 E} \left[-\frac{dS}{S^n} \right]$$

Integrating from the initial stress S_1 to the residual stress S

$$t = \frac{S_0^n}{r_0 E} \left[-\int_{S_1}^S \frac{dS}{S^n} \right] = \frac{S_0^n [1 - (S/S_1)^{n-1}]}{(n-1)r_0 E S^{n-1}} \dots \dots [3]$$

³ "The Resistance to Relaxation of Metals at High Temperature," by Ernest L. Robinson, Trans. ASME, vol. 61, 1939, pp. 543-554.

⁴ "High-Temperature Bolting Materials," by Ernest L. Robinson, Proceedings of the ASTM, vol. 48, 1948, pp. 214-235.

$$\left(\frac{S}{S_1} \right)^{n-1} = \frac{1}{(n-1)r_0 E t \frac{S_1^{n-1}}{S_0^n} + 1} \dots \dots [4]$$

Formula [3] shows the time that the bolt will stay tighter than some particular working stress level below which it is undesirable to go. Formula [4] shows the stress to which the bolt will relax at the end of any particular period of time. The difference between the initial stress and the residual stress divided by the modulus of elasticity measures the total amount of plastic extension in the bolt (see Formula [2]). This is shown by the inverse scale at the right of Fig. 1. This will form a convenient basis of comparison for more elaborate applications.

F. P. COFFIN'S SPRING-LOADED CREEP FURNACE

More than 25 years ago, when creep testing was not so far advanced as it now is, F. Parkman Coffin⁵ used long soft springs instead of weights to load the levers of one of his creep-test furnaces shown diagrammatically in Fig. 2.

The way this machine worked may be analyzed as follows: Note that the take-up on the loading wheel at the left is

$$\frac{S_1}{E} L + \frac{S_1 A}{K_1 K_2^2}$$

The first term represents the elasticity of the specimen itself while the second term is the follow-up elasticity of the long, soft spring. It is convenient to define the extra elasticity of the arrangement, relative to that of the creep bar, as β so that

$$\beta = \frac{\frac{A}{K_1 K_2^2}}{\frac{L}{E}} = \frac{AE}{LK_1 K_2^2}$$

After a period of time at any instant t the stress in the specimen is S . The total extension in the machine is always the same as the take-up on the loading wheel so that

$$\frac{S_1}{E} L + \frac{S_1 A}{K_1 K_2^2} = \frac{S}{E} L + \frac{SA}{K_1 K_2^2} + pL$$

$$\text{from which } S_1 - S = \frac{E}{(1 + \beta)} p \dots \dots [5]$$

Differentiating with respect to time gives the elastic requirement that

$$-\frac{dS}{dt} = \frac{E}{(1 + \beta)} \frac{dp}{dt}$$

But $dp/dt = r$ and from the creep properties

$$r = \frac{r_0}{S_0^n} S^n$$

$$\text{from which } dt = (1 + \beta) \frac{S_0^n}{r_0 E} \left[-\frac{dS}{S^n} \right]$$

Integrating from S_1 to S

$$t = \frac{(1 + \beta) S_0^n [1 - (S/S_1)^{n-1}]}{(n-1)r_0 E S^{n-1}} \dots \dots [6]$$

$$\left(\frac{S}{S_1} \right)^{n-1} = \frac{1}{(n-1)r_0 E t \frac{S_1^{n-1}}{S_0^n} + (1 + \beta) S_0^n} + 1 \dots \dots [7]$$

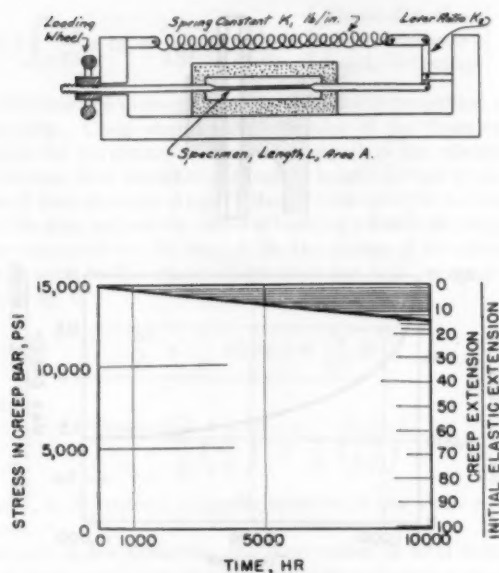


FIG. 2 COFFIN'S OLD SPRING-LOADED CREEP FURNACE WITH FOLLOW-UP ELASTICITY 100 TIMES THAT OF SPECIMEN

(2.25 per cent Cr-1 per cent Mo piping material at 1000 F
 $S_0 = 7500$, $n = 4$, $E = 20,000,000$, $(1 + \beta) = 101$)

Initial stress, 15,000; final stress, 12,740; total creep, 1.14 per cent. Note that amount of creep is about 20 times as much as would correspond with stress reduction if there were no follow-up elasticity. 10,000-hr creep is 15.22 times initial elastic extension.)

Formula [6] in comparison with [3] shows that the time required for relaxation from any particular initial stress to any particular final stress is prolonged in proportion to the follow-up elasticity. The follow-up elasticity of the soft spring prevents the reduction of stress due to creep which characterized the simple bolt in an unyielding flange. This is shown in Fig. 2 where the creep scale at the right now has to be based on $(1 + \beta)/E$ in accordance with Formula [5].

ORDINARY FLANGE WITHOUT GASKET

In an ordinary solid flange without a gasket, two influences occur which are different from the simple bolt in an unyielding flange. Because there is some small compressive stress in the flange itself, there is some additional elasticity besides that of the bolt itself although it is very much less than afforded by the long, soft springs of the old creep furnace. Also, because of the stress in the flange there is some additional creep besides what takes place in the bolt itself. Letting β represent the additional elasticity and ϕ the additional creep, the total take-up of the nuts is always equal to its initial value

$$(1 + \beta) \frac{S_1}{E} L = (1 + \beta) \frac{S}{E} L + (1 + \phi) p L$$

from which

$$S_1 - S = \frac{(1 + \phi)}{(1 + \beta)} E p \quad [8]$$

proceeding in the same manner as before to balance known elastic and creep properties

$$t = \frac{(1 + \beta) S_0^n [1 - (S/S_1)^{n-1}]}{(n-1) r_0 E (1 + \phi) S^{n-1}} \quad [9]$$

$$\left(\frac{S}{S_1}\right)^{n-1} = \frac{1}{(n-1)(1 + \phi) r_0 E t S_1^{n-1} + 1} \quad [10]$$

In order to cultivate some notion as to the importance of these several factors in a bolt and flange (all made of piping material) assume that the area of flange subject to compression is a circle whose diameter is twice that of the bolt as shown by the dotted lines in Fig. 3. The compressive stress on the net area, neglecting clearance, is $1/3$ the tension stress in the bolt. According to this assumption, $\beta = 1/3$ and $(1 + \beta) = 1.333$. The creep rate in the flange relative to that in the bolt is

$$\phi = \frac{\frac{r_0}{S_0^n} \left(\frac{1}{3} S\right)^n}{\frac{r_0}{S_0^n} S^n} = \left(\frac{1}{3}\right)^n$$

$$\text{For } n = 4 \quad \phi = \frac{1}{81} = 0.0123$$

Fig. 3 illustrates the behavior of such a bolt in a flange of the same material.

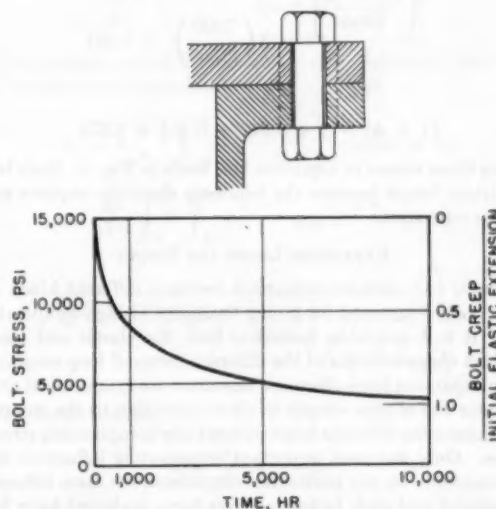


FIG. 3 BEHAVIOR OF AN ORDINARY BOLTED FLANGE WITHOUT A GASKET

(All made of 2.25 per cent Cr-1 per cent Mo piping steel; temperature, 1000°F
 $S_0 = 7500$, $E = 20,000,000$, $n = 4$, $(1 + \beta) = 1.333$, $(1 + \phi) = 1.0123$.
 Initial stress, 15,000; final stress, 4090; maximum creep, 0.072 per cent;
 10,000-hr creep is 0.96 times initial elastic extension.)

The formula for residual stress differs from that which describes a creeping test bar loaded by a cold spring which does not creep. In the new relationship, account is taken of both moderating influences even though the effect of the creep is of little consequence in this case.

BOLTS WITH COLLARS

A design feature used to keep bolts tighter for longer periods of time is to give them greater length by means of extension collars having a cross-sectional area equal to that of the bolt.⁴

⁴ "Progress in Design and Performance of Modern Large Steam Turbines for Generator Drive," by G. B. Warren, Trans. ASME, vol. 63, 1941, p. 71. (Describes stop valve bonnet bolts made longer by thick washers having a cross-sectional area equal to the bolts so as to provide three times as much flexibility.)

The additional elasticity of such bolts enables them more easily to absorb thermal shock without coming loose. More importantly for the present discussion, the bolt extensions and collars are likely to have 50 to 100 per cent greater creep strength by virtue of being 50 deg F cooler in temperature.

Formulas [8], [9], and [10] describe the behavior of such a bolted flange. The effect of the additional follow-up elasticity may be judged by assuming the bolt to be three times as long as in the ordinary flange, Fig. 3, and also making the same assumption that the compressive stress in the flange is $1/3$ as much. Thus

$$(1 + \beta) = 1 + 2 + \frac{1}{3} = 3.333$$

In order to evaluate the amounts of creep which occur in addition to that of the bolt where it goes through the flange, the extra creep in the flange is assumed to be $(1/3)^n = 0.0123$ as in the ordinary flange. However, in the extensions of the bolt the temperature is likely to be 50 deg F colder so that in that region S_0 is 11,500 instead of 7500. Consequently, the additional creep in that region relative to the bolt where it goes through the flange is

$$2 \left[\frac{r_0 S_0^n}{7500} \right] = 2 \left(\frac{7500}{11500} \right)^4 = 0.361$$

$$(1 + \phi) = 1 + 0.012 + 0.361 = 1.373$$

Putting these values in Equation [10] leads to Fig. 4. Such bolts stay tighter longer because the follow-up elasticity requires more creep for relaxation.

EXPANSION LOOPS AND PIPING

In order to facilitate comparison between different kinds and sizes of loops employed for giving flexibility to high-temperature piping, it is desirable to formulate both the elastic and plastic deflection characteristics of the different forms of loop employed. All the expansion loops discussed hereafter are symmetrical about the center run of pipe simply to direct attention to the comparisons between the different loops without any complicating circumstances. Only the most important contributing influences have been counted. In any particular comparison, the same influences are included and such factors as have been neglected have been neglected on both sides of the comparison. Thus the most important bending deflection is contributed by the curvature of the offset straight runs, whether elastic or plastic. However, the contributions due to the bending of the lengths at right angles to the main line of pipe are included. On the other hand, the refinement of precise consideration of behavior of quarter turns has not been considered in any of the comparisons. The same remarks apply equally to torsion loops where the bending of the connections at right angles to the main line of piping has been included but with no special consideration for the quarter turns.

FORMULAS FOR DEFLECTION DUE TO BENDING, WHETHER ELASTIC OR PLASTIC

Formulas for deflection due to bending are based upon deter-

⁶ ASTM Special Technical Publication No. 151, compiled by and issued under the auspices of the Data and Publications Panel of the ASME-ASTM Joint Committee on the Effect of Temperature on the Properties of Metals, gives on p. 90 data on the creep strength of 2.25 per cent Cr-1 per cent Mo steel plotted against temperature for two nominal rates from which the value of n may be determined. This is the source of the properties used in this paper.

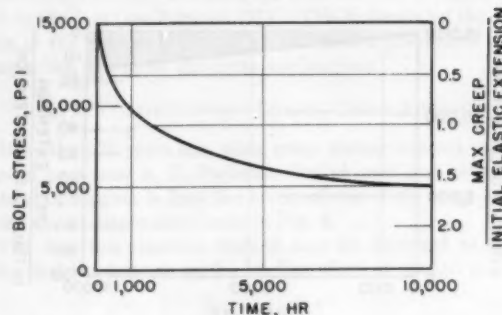
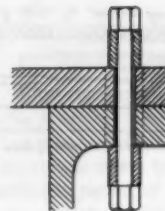


FIG. 4 BEHAVIOR OF BOLT WITH EXTENSION COLLARS IN FLANGE WITHOUT GASKET

(All made of 2.25 per cent Cr-1 per cent Mo piping steel. Flange and primary length of bolt 1000 F but extensions 50 deg F cooler $S_0 = 7500$, $n = 4$, $E = 20,000,000$, $(1 + \beta) = 3.333$, $(1 + \phi) = 1.373$. Initial stress, 15,000; final stress, 4980; maximum creep, 0.122 per cent; 10,000-hr creep is 1.62 times initial elastic extension.)

mination of curvature and the geometry of the length bent. The extension of the extreme fiber per unit of length divided by its distance from the neutral axis is the curvature C .

For elastic bending, if R is the outside radius of the pipe

$$C = \frac{S}{ER} \quad \text{and} \quad S = \frac{FlR}{I} \quad \text{so that} \quad C = \frac{Fl}{EI}$$

The contribution to total deflection of any particular length of pipe is found by multiplying its curvature by its length and its offset distance from the main line of pipe

$$D = Cal = \frac{S}{ER} al = \frac{Fl}{EI} al$$

In a simple symmetrical expansion loop such as shown in Fig. 5,

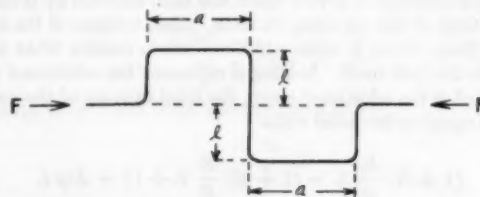


FIG. 5 SIMPLE SYMMETRICAL BENDING EXPANSION LOOP

there are two lengths a offset by a distance l . The elastic deflection due to these two lengths is

$$D = 2 \frac{Fl}{EI} al = 2 \frac{Fal^2}{EI} = 2 \frac{S}{ER} al \left\{ \begin{array}{l} \text{Elastic deflection} \\ \text{due to parallel} \\ \text{lengths in bending} \end{array} \right\} \dots [11]$$

There are four lengths l which are in the nature of cantilevers. The elastic deflection due to these four lengths is

$$D = 4 \int_0^l \frac{Fx}{EI} x dx = \frac{4Fl^3}{3EI} = \frac{4S}{3ER} \left\{ \begin{array}{l} \text{Elastic deflection} \\ \text{due to crosswise} \\ \text{lengths in bending} \end{array} \right\} \quad [12]$$

Deflection due to creep increases with time in proportion to the creep rate. Creep changes the extension of the fibers and so changes the curvature. Curvature is equal to the extension of the extreme fiber in inches per inch of length divided by its distance R from the neutral axis. (Here R stands for the half-diameter of the pipe and not the radius of bending which is the reciprocal of the curvature by definition.) So the change of curvature per hour is equal to the change of extension per hour, or creep rate, divided by R

$$\frac{dC}{dt} = \frac{r}{R} \text{ where } r = \frac{r_0}{S_0^n} S^n$$

so that

$$\frac{dC}{dt} = \frac{r_0}{R} \left(\frac{S}{S_0} \right)^n = \frac{r_0}{R} \left(\frac{FlR}{S_0 I_c} \right)^n$$

where I_c is the moment of inertia suitable to use when creep is present.

As with elastic deflection, the contribution to total deflection of any particular length of pipe is found by multiplying its curvature by its length and by its offset distance. This is also true for the rates per hour.

In the two lengths a

$$\begin{aligned} \frac{dD}{dt} &= 2 \frac{dC}{dt} a l = 2 \frac{r_0}{R} \left(\frac{S}{S_0} \right)^n a l \\ &= 2 \frac{r_0}{R} \left(\frac{FlR}{S_0 I_c} \right)^n a l \left\{ \begin{array}{l} \text{Deflection creep rate} \\ \text{due to parallel} \\ \text{lengths in bending} \end{array} \right\} \quad [13] \end{aligned}$$

where dD/dt is the increase of deflection in inches per hour.

Proceeding in a similar manner with reference to the four cantilever lengths l , the deflection rate is

$$\frac{dD}{dt} = 4 \int_0^l \frac{r_x}{R} x dx \text{ where } r_x = \frac{r_0}{S_0^n} S_x^n \text{ and } S_x = \frac{FxR}{I_c}$$

and x is the distance to the center line of the reactions

$$\begin{aligned} \frac{dD}{dt} &= 4 \int_0^l \frac{r_0}{RS_0^n} \left(\frac{FxR}{I_c} \right)^n x dx = 4 \frac{r_0 l^2}{(n+2)R} \left(\frac{FlR}{S_0 I_c} \right)^n \\ &= 4 \frac{r_0 l^2}{(n+2)R} \left(\frac{S}{S_0} \right)^n \left\{ \begin{array}{l} \text{Deflection creep rate} \\ \text{due to crosswise} \\ \text{lengths in bending} \end{array} \right\} \quad [14] \end{aligned}$$

It may be noted that the elastic deflections are in inches rather than per unit of length as in the case of the bolting examples. Similarly, the plastic deflection rates are in inches per hour instead of inches per inch of length per hour used with bolts. Each of these formulas for the deflection or its rate of increase with time is expressed both in terms of the over-all reaction F and in terms of the maximum stress S .

DETERMINATION OF MOMENT OF INERTIA FOR PLASTIC BENDING

Because the distribution of stress over the cross section of a pipe in bending differs under creep conditions from what it is when the material is elastic, a different and somewhat larger moment of inertia should be used to recognize the fact that the extreme fibers drop stress due to their susceptibility to creep while the fibers nearer to the neutral axis pick up a greater share of the moment to be carried.

Conditions of symmetry seem to justify the assumption that

plane sections stay plane as with elastic bending, but the creep stress has to accommodate itself to this geometry.⁷ Under elastic conditions the moment carried by a cross section of pipe is given by the familiar formula

$$M = \frac{I}{R} S$$

where S is the stress in the extreme fiber at a distance R from the neutral axis. It will be convenient to determine the value of

$$\frac{I_c}{I} = \frac{MR}{SI}$$

where I_c is the moment of inertia suitable to use under the conditions of plastic bending.

For a pipe whose outer radius is R and whose inner radius is R_i let $S_{r\alpha}$ be the stress at any distance $r \sin \alpha$ from the neutral axis. For this particular analysis r stands for radius instead of creep rate. But the creep rate is proportional to the distance from the neutral axis. Thus the ratio of the creep rate at distance $r \sin \alpha$ from the neutral axis to the rate at distance R , which is the outside surface where the stress has its maximum value S , is as follows

$$\frac{r \sin \alpha}{R} = \frac{S_{r\alpha}^n}{S^n} \quad S_{r\alpha} = S \left[\frac{r \sin \alpha}{R} \right]^{\frac{1}{n}}$$

The moment carried by the entire cross section is four times that carried by each quadrant

$$\begin{aligned} M &= 4 \int_0^{\frac{\pi}{2}} \int_{R_i}^R S \left[\frac{r \sin \alpha}{R} \right]^{\frac{1}{n}} r \sin \alpha r dr d\alpha \\ &= \frac{4S}{R^{1/n}} \int_0^{\frac{\pi}{2}} \int_{R_i}^R r^{2+\frac{1}{n}} [\sin \alpha]^{1+\frac{1}{n}} dr d\alpha \\ &= \frac{2\sqrt{\pi} S}{3+\frac{1}{n}} \left[R^3 - \frac{R_i^{3+\frac{1}{n}}}{R^{1/n}} \right] \frac{\Gamma\left(1+\frac{1}{2n}\right)}{\Gamma\left(1.5+\frac{1}{2n}\right)} \end{aligned}$$

For elastic bending

$$I = \frac{\pi}{4} (R^4 - R_i^4)$$

so

$$\frac{I_c}{I} = \frac{MR}{SI}$$

$$= \frac{8}{\sqrt{\pi} \left(3 + \frac{1}{n} \right)} \left[\frac{1 - (R_i/R)^{3+\frac{1}{n}}}{1 - (R_i/R)^4} \right] \frac{\Gamma\left(1+\frac{1}{2n}\right)}{\Gamma\left(1.5+\frac{1}{2n}\right)} \dots [15]$$

Logarithms of these gamma functions may be found in the back of Peirce's Table of Integrals. Fig. 6 shows that they never differ greatly from unity. Fig. 7 shows values of I_c/I plotted against the radius ratio for $n = 4$ and $n = 8$.

⁷ In this analysis of creep in bending, the author follows the same line of reasoning as used by Professor Marin in his book "Mechanical Properties of Materials and Design," McGraw-Hill Book Company, Inc., New York, N. Y., 1942, chapter 6. This book is out of print but the subject is also covered in Professor Marin's more recent book on "Engineering Materials," Prentice Hall, Inc., New York, N. Y., 1952.

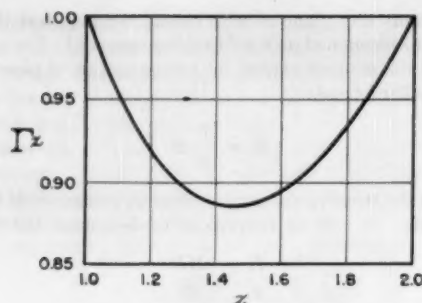


FIG. 6 THE GAMMA FUNCTION
(From Peirce's "Table of Integrals.")

It is a fair question how soon the pipe adjusts itself under creep conditions to the plastic stress distribution. It must start out with elastic distribution but this must disappear within a few days. As the author mentioned before, the relationships assumed are characteristic of long-time performance, that is, behavior beyond 1000 hr. It seems reasonable to suppose that stress distribution will be plastic after the first week or two.

SINGLE EXPANSION LOOP IN BENDING

The thermal expansion of the entire line is $\epsilon L \Delta T$ and when the line is assembled stress-free this is equal to the sum of the elastic deflections on coming up to temperature. For a loop like Fig. 5 from Equations [11] and [12]

$$\epsilon L \Delta T = 2 \frac{F_1 a l^3}{EI} + 4 \frac{F_1 l^3}{3EI}$$

This fixes the initial reactions just as the take-up of the nuts fixes the initial stress in a bolt. Thereafter the sum of the residual elastic deflections and the accumulated plastic deflection here represented by D , remains constant at this value.

$$2 \frac{F_1 a l^3}{EI} + 4 \frac{F_1 l^3}{3EI} = 2 \frac{F a l^3}{EI_c} + 4 \frac{F l^3}{3EI_c} + D$$

Differentiating with respect to time

$$\frac{dD}{dt} = - \left[\frac{2al^3}{EI_c} + \frac{4l^3}{3EI_c} \right] \frac{dF}{dt}$$

Note that the maximum stress is in the a lengths and equal to

$$S = \frac{F l R}{I_c} \quad F = \frac{S I_c}{l R}$$

Differentiating with respect to time

$$\frac{dF}{dt} = \frac{I_c}{l R} \frac{dS}{dt}$$

Thus

$$\frac{dD}{dt} = - \left[\frac{2al}{ER} + \frac{4l^2}{3ER} \right] \frac{dS}{dt}$$

The terms in the bracket specifically represent the increase of deflection per unit of stress in the worst location.

Note also that β is the ratio of the additional elasticity to that due to the maximum stress in the a lengths so that

$$\beta = \frac{4l^2}{3ER} = \frac{2l}{3a}$$

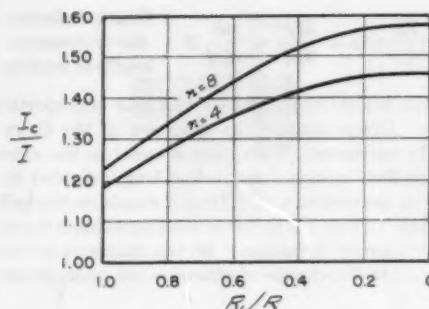


FIG. 7 RATIO OF MOMENT OF INERTIA FOR CREEP I_c TO I FOR ORDINARY ELASTIC BENDING PLOTTED FOR DIFFERENT THICKNESSES OF PIPE

Thus the elastic relationship between deflection and stress also may be written

$$\frac{dD}{dt} = - (1 + \beta) \left[\frac{2al}{ER} \right] \frac{dS}{dt}$$

The change of deflection with time is, of course, fixed by the creep properties, Equations [13] and [14]

$$\begin{aligned} \frac{dD}{dt} &= \left[2 \frac{r_0 a l}{R} + 4 \frac{r_0 l^3}{(n+2)R} \right] \frac{S^n}{S_0^n} \\ &= \left[\frac{2 r_0 l}{R} \left(a + \frac{2l}{(n+2)} \right) \right] \frac{S^n}{S_0^n} \end{aligned}$$

The terms in the bracket specifically represent the increase of deflection in inches per hour when the stress in the worst location is passing through the S_0 level.

Note also that ϕ is the ratio of the additional creep to that due to the maximum stress so that

$$\phi = \frac{4r_0 l^3}{(n+2)R} = \frac{2l}{(n+2)a}$$

Thus the time derivative of deflection also may be written

$$\frac{dD}{dt} = (1 + \phi) \left[\frac{2r_0 a l}{R} \right] \frac{S^n}{S_0^n}$$

Equating these two expressions for dD/dt

$$- \frac{2l}{ER} \left[a + \frac{2l}{3} \right] \frac{dS}{dt} = \frac{2r_0 l}{R} \left[a + \frac{2l}{(n+2)} \right] \frac{S^n}{S_0^n}$$

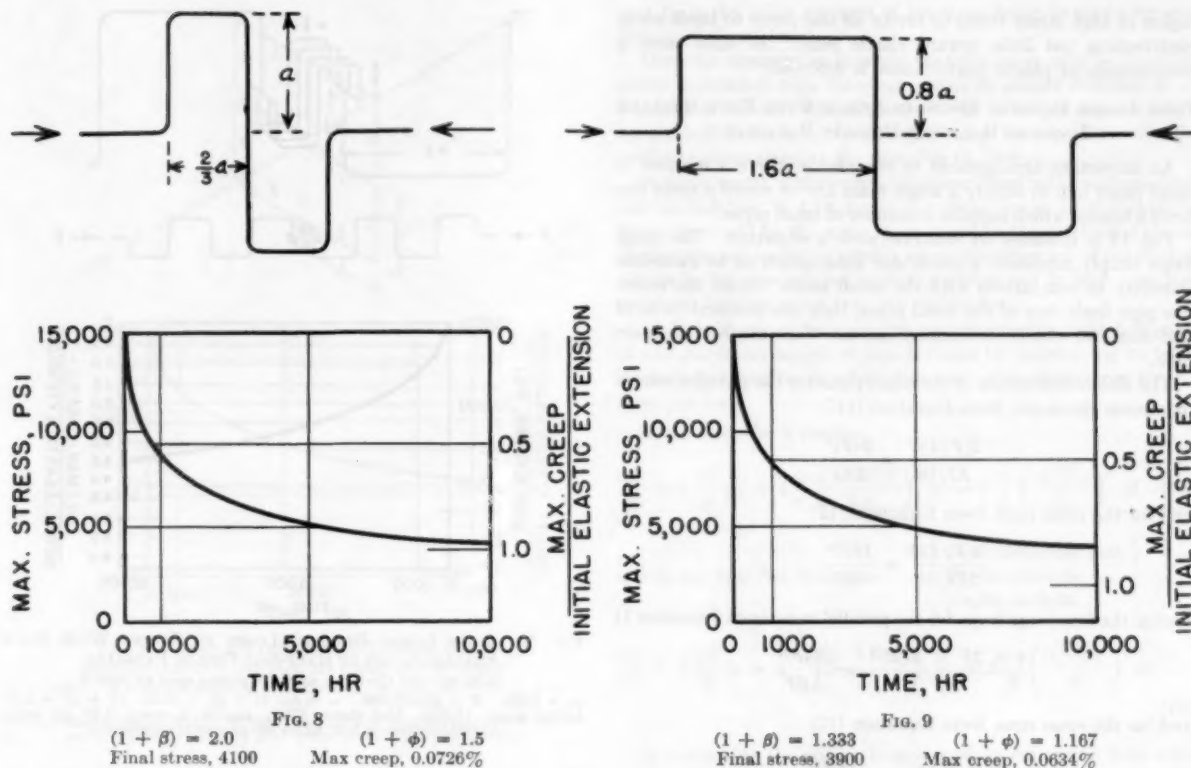
or alternatively

$$- (1 + \beta) \left[\frac{2al}{ER} \right] \frac{dS}{dt} = (1 + \phi) \left[\frac{2r_0 a l}{R} \right] \frac{S^n}{S_0^n}$$

From which

$$\begin{aligned} dt &= \left[\frac{1}{r_0 E} \left(\frac{1 + \frac{2l}{3a}}{1 + \frac{2l}{(n+2)a}} \right) \right] \left(- \frac{dS}{S^n} \right) S_0^n \\ &= \left[\frac{(1 + \beta)}{r_0 E (1 + \phi)} \right] \left(- \frac{dS}{S^n} \right) S_0^n \end{aligned}$$

which leads directly to Equations [9] and [10]. The expression in the bracket is the time for unit stress relief when the worst stress



FIGS. 8-9 SIMPLE BENDING EXPANSION LOOPS

(Both made of 2.25 per cent Cr-1 per cent Mo piping steel; temperature, 1000 F. $S_0 = 7500$, $E = 20,000,000$, $n = 4$, initial stress = 15,000.)
 (These two loops have the same elasticity and the same reactions, but the maximum 10,000-hr creep in the shorter wider loop of Fig. 8 is 0.97 times the initial elastic extension as compared with 0.85 for the longer narrower loop of Fig. 9.)

is equal to the nominal creep strength. Thus the procedure for analyzing a piping loop during "self-springing" is the same as for determining the residual stress in a bolt with both follow-up elasticity and extra creep.

Figs. 8 and 9 show two such expansion loops with equal elastic deflections for equal reactions. However, the loop shown in Fig. 8 depends on the cantilever lengths for half its elasticity. These parts add 100 per cent follow-up elasticity to the deflection provided by the offset subject to creep. On the other hand, the loop shown in Fig. 9 derives only a quarter of its elasticity from the low-stressed cantilever portions which add only a third follow-up. As a consequence the loop shown in Fig. 8 requires somewhat more creep for relaxation than the loop in Fig. 9.

The important thing to note about these simple loops is that they do not provoke any great amounts of creep concentration. In this respect they are similar to the ordinary bolt in an ordinary flange without gasket.

ONE LARGE LOOP WITH SEVERAL SMALLER LOOPS IN BENDING

An arrangement which is more likely to be accompanied by localization of creep is where there are a series of loops, one of which extends beyond the others. Such a case is illustrated in Fig. 10. It may be taken as typical of a combination of low-stressed elasticity with one particular region subjected to the same reactions but more highly stressed.

Although the stress-relief diagram looks very similar to those already considered, it should be noted that the 10,000-hr creep in the extremes of the one big loop is two thirds more than the total initial elastic extension. This is a feature of high-temperature piping design to avoid wherever possible because the one small

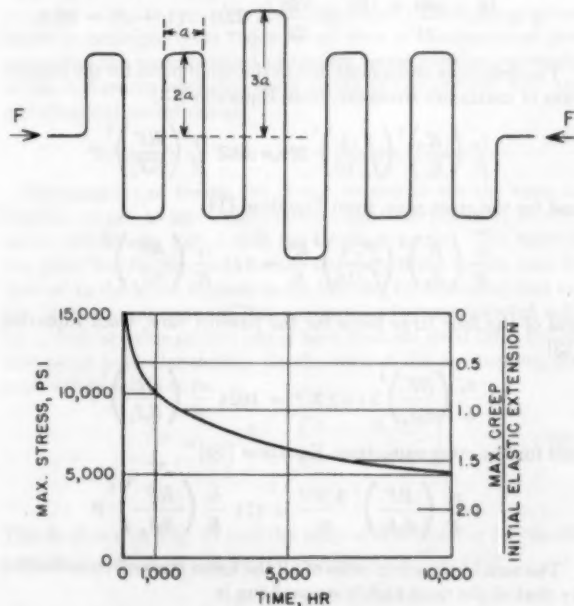


FIG. 10 ONE WIDER EXPANSION LOOP WITH FOUR NOT SO WIDE
 (2.25 per cent Cr-1 per cent Mo piping steel at 1000 F)

$S_0 = 7500$, $n = 4$, $E = 20,000,000$, $(1 + \beta) = 7.15$, $(1 + \phi) = 2.876$
 Initial stress, 15,000; final stress, 5010; maximum creep, 0.124 per cent;
 10,000-hr creep is 1.66 times initial elastic extension.)

region of high stress tends to invite all the creep to itself while contributing but little toward elastic relief. In such cases a computation of plastic performance is desirable.

FOUR LARGE BENDING LOOPS IN SERIES WITH FOUR SMALLER LOOPS OF HALF-SIZE PIPES IN PARALLEL

An important arrangement to consider is where a number of small pipes join to supply a single main line or where a main line feeds a header which supplies a number of small pipes.

Fig. 11 is intended to illustrate such a situation. The large loops simply represent a particular assumption as to main-line elasticity in comparison with the small loops. Since the main-line pipe feeds four of the small pipes, they are assumed to be of half-diameter with cross-sectional inertias $1/16$ as much as the main pipe.

The elastic deflections of the small pipes for the parallel runs of maximum stress are, from Equation [11]

$$\frac{2(F/4)l^3}{EI/16} = \frac{24Fl^3}{3EI}$$

and for the cross runs, from Equation [12]

$$\frac{4(F/4)l^3}{3EI/16} = \frac{16Fl^3}{3EI}$$

and of the four large loops for the parallel runs, from Equation 11

$$\frac{4 \times 2F \times 4l(2l)^3}{EI} = \frac{384Fl^3}{3EI}$$

and for the cross runs, from Equation [12]

$$\frac{4 \times 4F(2l)^3}{3EI} = \frac{128Fl^3}{3EI}$$

The sum of the elastic deflections of all the lesser stressed runs divided by that of the most highly stressed run is

$$\beta = \frac{16 + 384 + 128}{24} = \frac{528}{24} = 22.0 \quad (1 + \beta) = 23.0$$

The deflection creep rates due to the small pipes for the parallel runs of maximum stress are, from Equation [13]

$$\frac{\tau_0}{R} \left(\frac{R}{S_0} \right)^4 \left(\frac{F/4}{I_0/16} \right)^4 2l^3 = 512 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 l^3$$

and for the cross runs, from Equation [14]

$$\frac{\tau_0}{R} \left(\frac{R}{S_0} \right)^4 \left(\frac{F/4}{I_0/16} \right)^4 \frac{4l^3}{6} = 171 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 l^3$$

and of the four large loops for the parallel runs, from Equation [13]

$$4 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 2(4l)(2l)^3 = 1024 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 l^3$$

and for the cross runs, from Equation [14]

$$4 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 \frac{4(2l)^3}{6} = 171 \frac{\tau_0}{R} \left(\frac{RF}{S_0 I_0} \right)^4 l^3$$

The sum of the creep rates of all the lesser stressed runs divided by that of the most highly stressed run is

$$\phi = \frac{171 + 1024 + 171}{512} = \frac{1366}{512} = 2.67 \quad (1 + \phi) = 3.67$$

The curves in Fig. 11 are based on these figures and show that

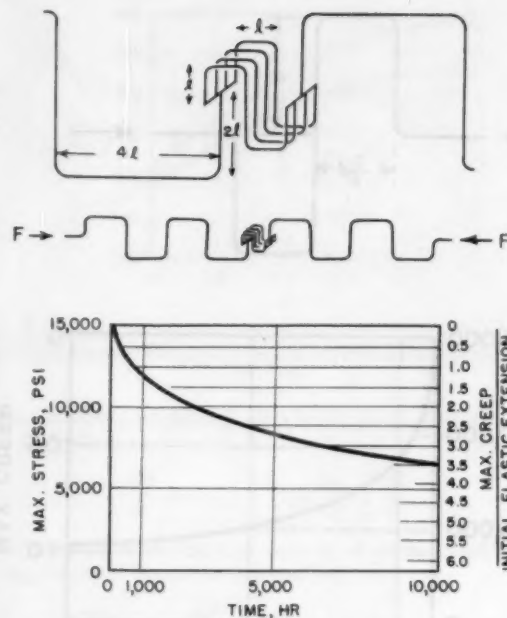


FIG. 11 FOUR LARGE BENDING LOOPS IN SERIES WITH FOUR SMALLER LOOPS OF HALF-SIZE PIPE IN PARALLEL
(2.25 per cent Cr-1 per cent Mo piping steel at 1000 F
 $S_0 = 7500$, $E = 20,000,000$, $n = 4$, $(1 + \beta) = 23.0$, $(1 + \phi) = 3.67$
Initial stress, 15,000; final stress, 6700; maximum creep, 0.26 per cent;
10,000-hr creep is 3.47 times initial elastic extension.)

the creep in the outer fibers of the small pipes will be about $1/4$ per cent at the first 10,000 hr which is about 3.5 times the initial elastic extension. When employing this kind of construction, analysis of creep accumulation is desirable.

DEFLECTION DUE TO TORSION WHETHER ELASTIC OR PLASTIC

Formulas for deflection due to torsion are based upon determination of twist and the geometry of the length twisted. The distortion of the extreme fiber divided by its distance from the neutral axis is the twist per unit length T .

For elastic torsion

$$T = \frac{S_T}{GR} \quad S_T = \frac{FlR}{I_p} \quad \text{or} \quad \frac{FaR}{I_p} \quad T = \frac{Fl}{GI_p} \quad \text{or} \quad \frac{Fa}{GI_p}$$

The contribution to total deflection of any particular length of pipe is found by multiplying its length by its twist and its offset distance from the main line of pipe

$$D = Tal = \frac{S_T}{GR} al = \frac{Fl}{GI_p} al \quad \text{or} \quad \frac{Fa}{GI_p} al$$

In a simple symmetrical expansion loop such as shown in Fig 12, the elastic deflection due to twisting in the No. 2 lengths amounts to

$$D = 4 \frac{Fl}{GI_p} al = 4 \frac{Fal^2}{GI_p} \dots \dots \dots [16]$$

and in the No. 3 lengths

$$\left\{ \begin{array}{l} \text{Elastic Deflection} \\ \text{due to crosswise} \\ \text{lengths in twist} \end{array} \right\}$$

$$D = 4 \frac{Fa}{GI_p} al = 4 \frac{Fa^2 l}{GI_p} \dots \dots \dots [17]$$

Because bending occurs as well as twisting, it also must be

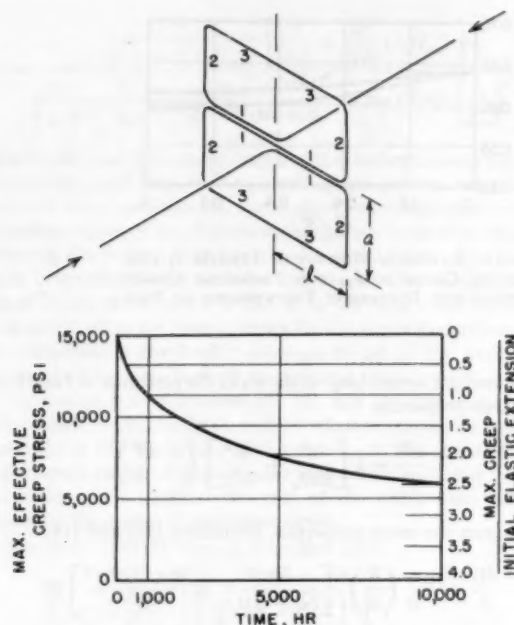


FIG. 12 SIMPLE SYMMETRICAL TORSION LOOPS
(2.25 per cent Cr-1 per cent Mo piping steel at 1000 F
 $S_0 = 7500$, $E = 20,000,000$, $G = 8,000,000$, $n = 4$, $(1 + \beta)/(1 + \phi) = 4.11$
Initial effective creep stress, 15,000; final, 5890; maximum creep, 0.187 per cent; 10,000-hr creep is 2.50 times initial elastic extension.)

recognized. There are four No. 1 lengths in the plane of the reactions in pure bending for which the deflections already have been written, (Equation [12])

$$D = \frac{4Fl^3}{3EI}$$

Equal amounts of bending occur in the four No. 3 lengths

$$D = \frac{4Fl^3}{3EI}$$

Also, similar bending occurs in the four No. 2 lengths

$$D = \frac{4Fa^3}{3EI}$$

The total elastic deflection of the loop is the sum of all these separate deflections

$$\begin{aligned} D &= 4 \frac{Fl^3}{3EI} + 4 \frac{Fa^3}{3EI} + 4 \frac{Fl^3}{3EI} + 4 \frac{Fa^3}{3EI} + 4 \frac{Fa^3l}{GI_p} \\ &= 4F \left[\frac{l^3 + a^3 + l^3}{3EI} + \frac{a^3 + a^3l}{GI_p} \right] \end{aligned}$$

For plastic twisting the distortion of the extreme fiber per unit of time is $\dot{\gamma}$ and the corresponding rate of change in the twist is

$$\frac{dT}{dt} = \frac{\dot{\gamma}}{R} \text{ where } \dot{\gamma} = \frac{r_0}{S_0} S^* \text{ and } S = \sqrt{3}S_T$$

so that

$$\frac{dT}{dt} = \frac{r_0}{R} \left(\frac{\sqrt{3}S_T}{S_0} \right)^n = \frac{r_0}{R} \left(\frac{\sqrt{3}FlR}{S_0 I_{pe}} \right)^n \text{ or } \frac{r_0}{R} \left(\frac{\sqrt{3}FaR}{S_0 I_{pe}} \right)^n$$

and I_{pe} is the polar moment of inertia suitable to use under conditions of plastic twist.

Here the assumption is made that the creep rate is quantitatively associated with the shear stress on planes disposed in an octahedral manner with reference to the principal stresses, commonly called the octahedral shear stress, rather than with the ordinary shear stress in the pipe due to torsion.

In other words, if the ordinary tension stress in the extreme fiber due to bending is 1.73 times the shear stress in the extreme fiber, there will be the same shear stress on the octahedral planes in the pipe subjected to torsion as on the other set of octahedral planes in the pipe subject to pure bending and in each case it is assumed that the creep rate will be the same.

As with elastic deflections, the contribution to total deflection of any particular length of pipe is found by multiplying its twist by its length and by its offset distance. This is also true for the rates per hour.

In the four No. 2 lengths

$$\frac{dD}{dt} = 4 \frac{dT}{dt} al = 4 \frac{r_0}{R} \left(\frac{\sqrt{3}FlR}{S_0 I_{pe}} \right)^n al = 4 \frac{r_0}{R} \left(\frac{\sqrt{3}S_T}{S_0} \right)^n al \quad [18]$$

and in the four No. 3 lengths

$$\left\{ \begin{array}{l} \text{Deflection creep rate} \\ \text{due to crosswise} \\ \text{lengths in twist} \end{array} \right\}$$

$$\frac{dD}{dt} = 4 \frac{dT}{dt} la = 4 \frac{r_0}{R} \left(\frac{\sqrt{3}FaR}{S_0 I_{pe}} \right)^n la = 4 \frac{r_0}{R} \left(\frac{\sqrt{3}S_T}{S_0} \right)^n la \quad [19]$$

In computing the plastic behavior of an expansion loop with both bending and torsional flexibility, the author recommends the simplest analysis that seems adequate.

For determining the reactions, it is assumed that torsional creep and bending creep act independently of each other and may be added directly, although the stresses resulting are to be compounded with due regard to their directions. These assumptions could be criticized from the point of view of mathematical precision but the recommended procedure seems justified especially as the influences neglected seem minor in comparison with the metallurgical uncertainties.

MOMENT OF INERTIA FOR PLASTIC TORSION

The moment of inertia for plastic torsion is not the same as what is proper to use under elastic conditions nor is it, as with elastic conditions, just double the bending inertia. The ratio of the polar inertia for creep to the ordinary polar inertia may be derived in the same manner as for bending by assuming that radial lines stay straight. (This assumption is recommended only for a circular cross section but is here assumed good for a hollow pipe as for a circular shaft.) On the basis of this assumption, the ratio will be found to be

$$\frac{I_{pe}}{I_p} = \frac{4}{\left(3 + \frac{1}{n}\right)} \left[\frac{1 - (R_1/R)^3}{1 - (R_1/R)^4} \right] \dots \dots \dots [20]$$

This is shown in Fig. 13 and the ratio of the bending inertia for creep to the polar inertia for creep as shown in Fig. 14 is

$$\frac{I_c}{I_{pe}} = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(1 + \frac{1}{2n}\right)}{\Gamma\left(1.5 + \frac{1}{2n}\right)} \frac{1 - (R_1/R)^{3 + \frac{1}{n}}}{1 - (R_1/R)^3} \dots \dots \dots [21]$$

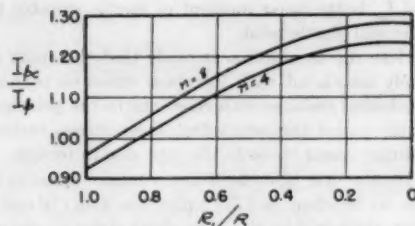


FIG. 13 RATIOS OF POLAR MOMENT OF INERTIA FOR CREEP I_{pe} TO I_p FOR ORDINARY ELASTIC TORSION PLOTTED FOR DIFFERENT THICKNESSES OF PIPE

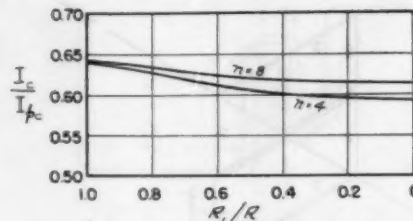


FIG. 14 RATIOS OF MOMENT OF INERTIA I_e FOR BENDING CREEP TO I_{pe} FOR TORSIONAL CREEP PLOTTED FOR DIFFERENT THICKNESSES OF PIPE

MEMORANDUM ABOUT THE "OCTAHEDRAL SHEAR STRESS"

Because the basic assumption is made that creep is a distortion of the material without dilation, it is useful to make note of the ordinary formulas which describe this behavior.

The shear stress on planes which are octahedral to the principal stresses is

$$S_{Oet} = \frac{1}{3} \sqrt{(S_1 - S_2)^2 + (S_1 - S_3)^2 + (S_2 - S_3)^2}$$

In an ordinary tensile bar S_2 and S_3 are both zero so that

$$S_{Oet} = \frac{1}{3} \sqrt{2S_1^2} = \frac{\sqrt{2}}{3} S_1$$

All ordinary creep data are expressed in terms of

$$S = \frac{3}{\sqrt{2}} S_{Oet} = 2.120 S_{Oet}$$

Therefore, in determining creep rates under compound stresses, it is necessary to multiply the octahedral shear stress by this ratio if ordinary creep data are to be used.

At the surface of an ordinary pipe if S_B is the longitudinal bending stress, S_T the torsional shear stress, and S_P the transverse tension due to pressure (bursting stress), the octahedral shear stress is

$$S_{Oet} = \frac{\sqrt{2}}{3} \sqrt{S_B^2 - S_B S_P + S_P^2 + 3S_T^2}$$

For determination of creep rates, this stress is to be multiplied by $3/\sqrt{2}$ so that

$$r = \frac{r_0}{S_0^n} S^n \text{ where } S = \sqrt{S_B^2 - S_B S_P + S_P^2 + 3S_T^2}$$

SINGLE EXPANSION LOOP IN TORSION

Fig. 12 illustrates a simple symmetrical expansion loop in which torsion is employed as well as bending to give flexibility to the line. When it is assembled without cold springing the initial reactions are fixed by the thermal expansion of the entire line. Thereafter the sum of the residual elastic deflections and the accumulated plastic deflection remains constant. From Equations [12], [16], and [17]

$$\begin{array}{ccccc} \text{Due to bending in} & & \text{Due to twist in} & & \\ \text{No. 1} & \text{No. 2} & \text{No. 3} & \text{No. 2} & \text{No. 3} \end{array}$$

$$D + 4F \left[\frac{l^3 + a^3 + l^3}{3EI_e} + \frac{al^3 + a^3l}{GI_{pe}} \right] = K$$

If the loop is a square loop as shown in the picture $a = l$ and from the elastic properties

$$\frac{dD}{dt} = 4 \left[\frac{3l^3}{3EI_e} + \frac{2l^3}{GI_{pe}} \right] \left(-\frac{dF}{dt} \right)$$

While from the creep properties, Equations [14] and [18]

$$\frac{dD}{dt} = 4 \frac{r_0}{R} \left(\frac{R}{S_0} \right)^n \left[\frac{3l^{n+2}}{(n+2)I_e^n} + \frac{2(\sqrt{3})^n l^{n+2}}{I_{pe}^n} \right] F^n$$

Because these two expressions are equal

$$dt = \frac{4 \left[\frac{l^3}{EI_e} + \frac{2l^3}{GI_{pe}} \right]}{4 \frac{r_0}{R} \left(\frac{R}{S_0} \right)^n \left[\frac{3l^{n+2}}{(n+2)I_e^n} + \frac{2(\sqrt{3})^n l^{n+2}}{I_{pe}^n} \right]} \left(-\frac{dF}{F^n} \right)$$

The foregoing relationship would be proper to use for direct determination of the decreasing reaction. However, for determining the extent of plastic distortion the corresponding expression in terms of stress is more useful. Being guided by the octahedral shear stress the maximum effective creep stress will be located where the No. 2 runs join the No. 3 runs and will be either

No. 2

No. 3

$$S = \sqrt{\left(\frac{FaR}{I_e} \right)^2 + 3 \left(\frac{FlR}{I_{pe}} \right)^2} \text{ or } S = \sqrt{\left(\frac{FlR}{I_e} \right)^2 + 3 \left(\frac{FaR}{I_{pe}} \right)^2}$$

$$= \frac{FlR}{I_e} \sqrt{(a/l)^2 + 3(I_e/I_{pe})^2} \text{ or } = \frac{FlR}{I_e} \sqrt{1 + 3(aI_e/I_{pe})^2}$$

whichever is larger.

For this particular square loop the maximum creep stress occurs each side of the elbow and is

$$S = \frac{FlR}{I_e} \sqrt{1 + 3(I_e/I_{pe})^2}$$

Reference to Fig. 14 will show that the value of this radical for pipes having a radius ratio $R_1/R = 0.75$ is 1.47 if $n = 4$ and 1.48 if $n = 8$.

Note that

$$F = \frac{SI_e}{lR \sqrt{1 + 3(I_e/I_{pe})^2}} \text{ and } dF = dS \frac{I_e}{lR \sqrt{1 + 3(I_e/I_{pe})^2}}$$

The time derivative may now be rewritten in terms of the effective creep stress instead of the end reaction

$$dt = \frac{\frac{4}{\sqrt{1+3(I_c/I_{pe})^2}} \left[\frac{l^2}{ER} + \frac{2l^2}{GR} (I_c/I_{pe}) \right]}{\frac{4}{\sqrt{1+3(I_c/I_{pe})^2}} r_0 \left[\frac{3l^2}{(n+2)} + 2(\sqrt{3})^n l^2 \left(\frac{I_c}{I_{pe}} \right)^n \right]} \times \left(-\frac{dS}{S^n} \right) S_0^n$$

Specifically, the two upper terms of this last expression represent the elastic deflections (due to bending and torsion, respectively) per unit creep stress in the worst elbow, while the two lower expressions represent the deflection changes per hour when the creep stress in the worst elbow has a value equal to S_0 . The over-all ratio is specifically the time for unit stress relief when the creep stress in the worst elbow is S_0 . It is the time constant for the loop and it gives the time in hours for the stress to decrease 1 psi. Its reciprocal is the hourly decrease in psi of the worst creep stress when it passes through the nominal stress level S_0 .

Comparing with Equations [8], [9], and [10] the numerator of this fraction is specifically $(1+\beta)/E$ which expresses the increased elasticity of the whole loop with reference to the region of maximum stress while the denominator is specifically $(1+\phi)r_0$ which expresses the increased creep rate of the whole loop with reference to the region of maximum stress.

The time constant may be simplified into

$$K = \frac{\left[1 + 2 \frac{E}{G} \frac{I_c}{I_{pe}} \right] \left[\sqrt{1 + 3 \left(\frac{I_c}{I_{pe}} \right)^2} \right]^{n-1}}{r_0 E \left[\frac{3}{(n+2)} + 2(\sqrt{3})^n \left(\frac{I_c}{I_{pe}} \right)^n \right]} = \frac{(1+\beta)}{r_0 E (1+\phi)}$$

and rewriting $dt = -KS_0^n dS/S^n$

$$\text{from which } \left(\frac{S}{S_0} \right)^{n-1} = \frac{1}{(n-1) \frac{t}{K} \frac{S_0^{n-1}}{S_0^n} + 1} \dots [22]$$

which is identical with Equation [10].

Fig. 15 gives the solution of this formula for the stated conditions. For the loop illustrated in Fig. 12 and assuming $R_1/R = 0.75$ for which $I_c/I_{pe} = 0.62$ and $E/G = 2.5$

$$\frac{(1+\beta)}{(1+\phi)} = 4.11 \quad \text{and} \quad K = 2.05 \text{ hr per psi}$$

and the maximum creep is 2.50 times the initial elastic extension.

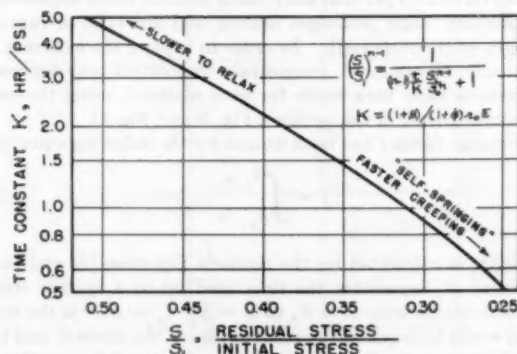


FIG. 15 TIME CONSTANT VERSUS 10,000-HR RESIDUAL-STRESS RATIO

2.25 per cent Cr-1 per cent Mo piping steel at 1000 F

$S_0 = 7500$, $E = 20,000,000$, $n = 4$

Initial stress, $S_1 = 15,000$, $t = 10,000$ hr

(Lest the catch words mislead, they refer to the pipe line as a whole. A self-springing line is likely to creep less in the worst place while a line which is slow to relax keeps on forcing creep in the worst place.)

It should be noted here that the same initial stress of 15,000 has been assumed as in the previous cases. Here it stands for an octahedral shear stress of

$$\frac{15,000}{3/\sqrt{2}} = 7080$$

The new wording for the flexibility section of the Code for Pressure Piping limits the combined expansion stress in a somewhat different manner to the maximum shear stress on the surface which is $\sqrt{S_B^2 + 4S_T^2}$ instead of $\sqrt{S_B^2 + 3S_T^2}$

CONCLUDING REMARKS

The foregoing discussion and examples give a selection of formulas for the determination of possible creep concentrations in high-temperature piping and illustrate some specific arrangements which are likely to be conducive to the concentration of plastic strain. Excessive plastic strain is undesirable, although the amount which might be tolerated varies greatly between different materials.

In the examples given by way of illustration, the assumption is made that the pipe is assembled in a stress-free condition when cold, that is, without cold springing. The term self-springing is another way of saying that relaxation occurs. When most of the relaxation or self-springing is furnished in a local region undesirably large permanent plastic strains or creep are possible. This is one reason why cold springing is desirable and it should preferably be done in such a manner that the pipe will be practically stress-free at temperature. When this is done, the stresses due to thermal expansion will be present when the piping is cold and strong and not subject to creep.⁸

The two most important assumptions are reviewed once more. The first is that creep goes with some power of the stress. However important, this rule may fall short of being a universal description of the initial stages of strain-hardening for all materials, but it is certainly a good comparative rule for any particular material in long-time service at high temperature. The author believes that comparisons between designs made on this basis are good comparisons.

The other basic assumption used for the evaluation of designs where both torsion and bending are present is the assumption that the progressive using up of the material goes with the octahedral shear stress. This is a nondirectional stress, except as it is associated with the directions of the principal stresses. In magnitude, it is fixed by the square root of the sum of the squares of the three maximum shear stresses and it is useful as a measure of the energy of distortion, that is, the energy which goes to change shape without changing volume. This is a characteristic of creep strain as distinguished from elastic strain which is accompanied by dilation or change of volume.

Finally, in trying to decide if a structure is safe and useful, it is well to remember that fracture and failure come about through a

⁸ "Progress in Design and Performance of Modern Large Steam Turbines for Generator Drive," by G. B. Warren, Trans. ASME, vol. 63, 1941, p. 75. ("...whereas creep considerations alone indicated that cold straining of a piping system might not be necessary since 'creep would soon relieve it anyway,'... slow tests to rupture indicate that if much strain is necessitated by slow adjustments to some localized overstressed conditions, cracks might be started. This could be guarded against by cold-straining the piping system nearly to its hot condition.")

separation of one part from another and this is literally a change of shape which is technically called strain and that stress, as such, is just an intermediate means by which strain is caused.

Appendix

NOMENCLATURE

The following nomenclature is used in the paper:

- S = stress, psi
- S_0 = nominal creep strength, psi, for stated rate
- S_1 = initial stress, psi
- S_B = longitudinal stress due to bending, psi
- S_T = shear stress due to torque, psi
- E = modulus of elasticity, psi
- G = modulus of elasticity in shear, psi
- r = rate of creep, in. per in. per hr
- r_0 = nominal creep rate, 10^{-7} or $1/10,000,000$ in per in. per hr
- n = slope of log-log plot of tests r versus S
- t = time, hr
- F = reaction force at pipe ends, lb
- L = length, in.
- a = loop dimension, in.
- l = loop dimension, in.
- x = distance along pipe, in.
- r = distance from center of pipe, in.
- I = moment of inertia for elastic bending, in.⁴
- I_c = moment of inertia for bending creep, in.⁴
- I_p = polar moment of inertia for elastic torsion, in.⁴
- I_{pc} = polar moment of inertia for torsional creep, in.⁴
- R = radius of pipe, in.
- A = area of cross section, sq in.
- C = curvature, 1/in. (reciprocal of radius of curvature)
- T = twist, 1/in.
- D = deflection, in. (whether elastic or plastic)
- p = plastic extension or creep, in. per in.
- β = additional elasticity relative to part in question
- ϕ = additional creep relative to part in question
- K, K_1, K_2 = constants
- ϵ = coefficient of thermal expansion in. per in. per deg F
- ΔT = change of temperature, deg F

Discussion

A. R. C. MARKL.⁹ A truly scientific analysis of the stresses and strains created in a piping system under thermal expansion, or one taking advantage of the full extent of present-day theoretical knowledge, would present a task of forbidding complexity. To make it possible, with available engineering personnel, to obtain at least a rough appraisal of the flexibility of a reasonable proportion of the line mileage installed each year, the problem has to be greatly simplified. This is accomplished by the use of more or less arbitrary assumptions and approximate methods. The resulting mechanization of calculations leads to a disregard of certain phases which may assume importance in borderline cases.

Like the paper presented by J. E. Brock¹⁰ last year on the effect of supports, the author's paper serves to direct attention to an aspect which the inexperienced designer might be inclined to ignore completely, either because he does not know about it or because no explicit rules for taking account of it have been

established. No claim is advanced that either of these effects need be evaluated in each individual system, but they should receive consideration. The writer believes that a competent analyst will instinctively avoid trouble from either source, in the first instance, by using properly designed spring supports at critical locations; in the second, by installing anchors, guides, or stops, in such a way that small-size lines will not be unduly imposed upon by larger, lower-stressed portions of the system. However, there is no question that the necessary instinct or "feel" must be acquired, either in the school of hard knocks or, preferably, through having investigated at least a few cases by theoretical calculation.

The author's development provides as simple a means for estimating total creep at critical locations as probably can be devised, and thus makes it possible to compare different systems which are equally stressed for the same amount of expansion with respect to their proneness to creep concentration. As expected, a study of different configurations shows that systems will be the more favorable in this respect, the higher the average stress (taken over the entire length of the specific system) is in relation to the peak stress.

While the author uses his findings to lend support to his long-time advocacy of 100 per cent cold spring, he has not established any criterion by which the damage from initial overstressing of a line installed without the benefits of cold spring can be assessed. The fact that creep can be localized is indisputable, but many engineers, including the writer, have difficulty in visualizing failure under relaxation conditions in a line designed to the stress limits given in the Piping Code provided the material has good initial ductility, even though this may be greatly reduced under prolonged heating.

To clarify this point in his own mind, the writer has made two pairs of calculations directed at ascertaining to what extent initial overstressing of a line will reduce its service life. To obtain the maximum damage possible, the initial stress has been taken at the extreme value of the expansion stress permitted in the Code for Pressure Piping under the approved 1955 rules, this being $S_A = 1.25 (S_c + S_h)$ where S_c and S_h are the cold and hot allowable stresses (in this instance, as taken from the Power Piping Section); the final stress, after relaxation, has been set equal to S_h which is the pressure stress the Code sanctions for continuous operation at temperature.

Two materials have been assumed, the first, 2 1/4 per cent chrome-1 per cent moly steel, conforming to the author's choice, the second, carbon 1/2 per cent moly, being selected for its notoriously low ductility under prolonged heating and low ratio of rupture strength to creep strength. In order to explore the extremes of practically encountered creep-strain concentrations, duplicate calculations have been made for each material, using the configurations shown in the author's Fig. 9 and Fig. 11.

A damage factor f has been defined by the following equation

$$f = \int_{S_A}^{S_h} \frac{dt_u}{t_a} \dots \dots \dots [23]$$

Herein, t_a is calculated by the author's Equation [9] and each increment dt_u represents the time used up at a specific stress S , which relaxes from $S = S_A$ to $S = S_h$; $t_a = KS^m$ is the time which would be required to cause rupture at the stress S , and the materials constants K and m therein are obtained from a plot of log rupture time versus log stress-to-rupture. The life remaining after relaxation from S_A to S_h becomes $(1 - f)t_a$, and the total life $(1 - f)t_a + t_a$, where both t_a and t_u are now specific values applying to $S = S_h$. The relation between the total life for the initially overstressed condition and that where S_h is maintained continuously, expressed in per cent, then becomes

⁹ Research Consultant, Tube Turns, Louisville, Ky.

¹⁰ "Flexibility of Piping Systems Supported by Equally Spaced Rigid Hangers," by J. E. Brock, *Journal of Applied Mechanics*, Trans. ASME, vol. 76, 1954, pp. 11-18.

TABLE 1 TEST DATA AND RESULTS

	Metal temp 1000 F			
	Case no.		Case no.	
Pipe material.....	1	2	3	4
Hot modulus of elasticity E , psi ^a	23 × 10 ⁶		23 × 10 ⁶	
Allowable cold stress S_c , psi ^b	15000		13750	
Allowable hot stress S_h , psi ^b	7800		6250	
Allowable expansion stress range S_A , psi ^c	28500		25000	
Reference creep strength S_0 , psi.....	10000		10000	
Corresponding creep rate $\dot{\epsilon}_0$, (10 ⁻⁷ /hr) ^d	0.65 × 10 ⁻⁴		0.71 × 10 ⁻⁴	
Slope n of plot of $\log \dot{\epsilon}$ versus $\log S$ ^d	4.15		4.85	
Stress for rupture in 100,000 hr, psi ^d	14800		6800	
Factor k for time-rupture curve ^d	2.037 × 10 ¹⁴		2.767 × 10 ¹⁵	
Exponent m for time-rupture curve ^d	-9.426		-3.768	
Shape of system (author's Fig. No.).....	Fig. 9	Fig. 11	Fig. 9	Fig. 11
Corresponding creep factor (1 + ϕ).....	1.167	3.67	1.167	3.67
Corresponding elastic strain factor (1 + β).....	1.333	23.0	1.333	23.0
Ratio $C = (1 + \phi)/(1 + \beta)$	0.875	0.160	0.875	0.160
Results:				
Time t (hr) to relax to stress S_A	523	2872	1100	5994
Life τ (%) under relaxation from S_A to S_h in per cent of life for continuous operation at S_h ^f	98	88	97	82

^a "Elastic Constants and Coefficients of Thermal Expansion of Piping Materials Proposed for 1954 Code for Pressure Piping," by R. Michel, Trans. ASME, vol. 77, February, 1955, pp. 151-159.

^b From Supplement No. 1 to Code for Pressure Piping, ASA B31.1-1951, for material to ASTM A335, Gr. P22 and P1, respectively; all values are taken from Table 3a in the Power Piping Section, except that for carbon moly at 1000 F which, in the absence of power piping stresses, is taken from Table 21 in the Oil Piping Section.

^c From Report of Task Force of ASA Sectional Committee B31, revised Oct. 26, 1954, par. 621(c).

^d From "Digest of Steels," Fifth Edition, 1946, pp. 28 and 80, published by Timken Roller Bearing Company, Canton, Ohio.

^e By author's Equation [9], using $S_1 = S_A$ and $S = S_h$.

^f By writer's Equation [24].

$$\tau = 100[1 - f + t_u/t_a] \dots \dots \dots [24]$$

The pertinent data for the specific examples chosen and the final results obtained are given in Table 1. The relaxation curves for the four cases investigated and the stress-rupture curves for the two materials are plotted in Fig. 16, herewith. The former reveal that relaxation initially proceeds at a fast rate, a week's operation sufficing to bring the stresses down to one half their original levels; less than one or more than six months may be required for the stresses to decay to the values sanctioned by the Code for continuous operation, depending upon the follow-up elasticity of the system and the creep properties of the material.

Referring to the last line in Table 1, it is found that a reasonably balanced system, such as shown in Fig. 9 of the paper, relaxes with sufficient rapidity that its loss in useful life is only minor. On the other hand, a poorly designed layout with highly localized creep, exemplified by Fig. 11, may expend nearly one fifth its life in its preliminary adjustment. In each case, carbon moly has suffered about 50 per cent more loss in life than 2 1/4 per cent chrome-1 per cent molybdenum steel, which is partially due to its creep characteristics and partially due to the fact that the allowable stresses are set considerably closer to the rupture stress than is the case for 2 1/4 per cent chrome-1 per cent moly.

To forestall criticism, the writer would like to make it clear that he considers his analysis as furnishing no more than a first estimate. Within this limitation, however, it does serve to show that, while initial overstress has unfavorable consequences, these are not at all serious where a reasonably intelligent layout is

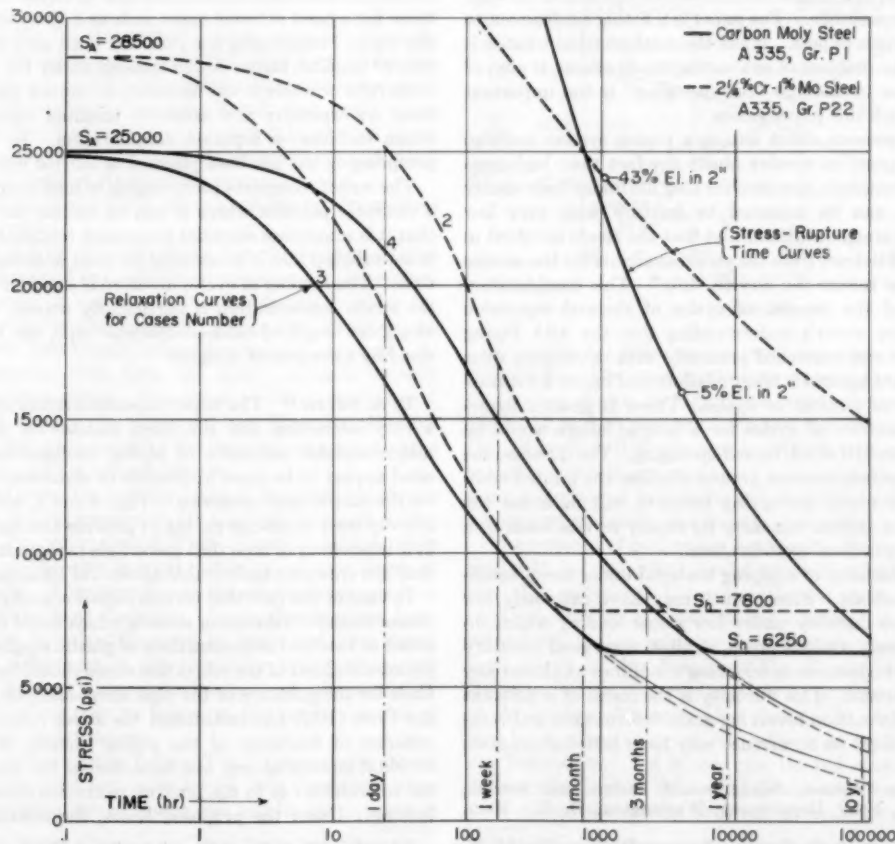


FIG. 16 RELAXATION AND STRESS-RUPTURE CURVES

used. There is no question that proper cold springing is beneficial in any event; but since it is often difficult to achieve, it would appear unreasonable and uneconomical to demand its application in every instance.

R. MICHEL.¹¹ In the past, designers of high-temperature piping have considered thermal expansion stresses as purely elastic, arriving at values in a manner similar to that used by a structural engineer in computing stresses in a bridge or other structure. Such methods do not take into account the fact that the pipe dimensions are permanently altered by creep at elevated temperatures. Because of this permanent growth of the metal, there is a reduction of thermal stress. This appears as a stress of opposite sign when the pipe is cooled.

A significant advance in pipe-design methods was made when cognizance was taken of the stress range from the cold to the hot condition and also the number of stress reversals. Stress concentrations are computed by multiplying stresses obtained in the usual manner by experimentally obtained stress-concentration factors.

The author has shown in his excellent paper that plastic growth, or creep, will be concentrated at points of high stress in certain pipe configurations and that these may assume dangerous proportions in certain instances. The paper is thus a significant contribution to the literature on high-temperature piping calculations and should become a tool of the high-temperature piping designer.

J. J. MURPHY.¹² The author is to be congratulated for taking the complex problem of creep relaxation of a high-temperature piping system and, in a simple and systematic mathematical way, reducing it to its essentials. The paper is a fitting supplement to the author's previous papers. While the mathematical treatise is most welcome, the problem of evaluating the ill effects, if any, of allowing the pipe to undergo "self-springing" is the important practical one which this paper raises.

The author's concern about letting a piping system undergo self-springing appears to revolve about the fact that high-temperature piping materials, designed for long life under their steady operating loads, can be expected to fracture with very low elongation under creep conditions and that the strain involved in self-springing will detract from the strain available for the service loadings and thus reduce the service "life." This consideration is independent of the number of cycles of thermal expansion loading. It is the writer's understanding that the ASA Piping Code Committee was concerned primarily with developing rules which would guard against a fatigue failure and insure a reasonable life in terms of number of cycles. There is general agreement that the number of cycles for a fatigue failure would be negligibly influenced, if at all, by self-springing. The question the author raises therefore revolves around whether the limited additional hot strain which self-spring involves will influence the length of time the system can carry its steady service loads to a degree requiring practical consideration.

The fracture ductility of a piping material under creep conditions is not a constant. Steels which may have extremely low long-time fracture ductility under low stress loading which induces predominately viscous creep, exhibit very good ductility under short-time high-stress overloading conditions at almost any stage prior to fracture. This ductility is the result of a different form of creep deformation involving grain deformation and rotation and crystal slip; its occurrence may have little influence on

the limiting fracture ductility under purely viscous creep conditions since a redistribution of internal stresses may mollify the accumulation of potential damage due to the purely viscous flow. The practical assessment of the problem is therefore extremely complex.

A piping system which has not been presprung can effect thermal stress relaxation by yielding first-stage creep and second-stage creep. Yielding is akin to hot-forming fabrication operations which are accepted without question and may not be involved depending on the initial stress level. First-stage creep, which occurs quickly, can probably be similarly regarded. The portion taking place as second-stage creep will detract from the service life but the bulk of it occurs over only a relatively short time compared to the service life and involves initial stresses likely to involve grain deformation and better inherent ductility during this period. Relaxation will be hastened by the existence of other longitudinal stresses in the piping such as weight stresses. It has generally been considered that the unit strains and relaxation time for uniform diameter lines of normal configurations will be sufficiently short so that any influence on life would be acceptably small. In a practical approach it must also be recognized that all high-temperature lines, whether presprung or not, are called upon to undergo plastic redistribution of stresses from time to time due to yielding, settlement, or maladjustment of supports, unequal heating or transient thermal effects, etc., and that the material used must have reasonable capacity to absorb such adjustments.

A vast number of high-temperature lines have been installed without prespringing and with no apparent ill effects. Occasional difficulties have been experienced due to strain concentration but these have been extreme cases such as a short length of reduced-size pipe. Prespringing has generally been used only where it was desired to utilize higher design stresses under the previous Piping Code rules or where it was necessary to control the initial hot reactions on expensive and sensitive terminal equipment, such as steam turbines—a separate consideration. In many cases the prespringing has not been planned or carried out effectively.

The writer recognizes prespringing of high-temperature lines as a desirable practice where it can be carried out effectively and that it is sometimes essential to protect terminal equipment, but is not satisfied that it is essential for most well-designed configurations. The number of cycles involved is a more vital factor unless the strain concentration is particularly severe; situations, such as a short length of reduced-diameter pipe, can be readily recognized by a competent designer.

R. A. SMYTH.¹³ The paper presents a rather difficult subject in a very interesting and simplified manner by clearly reviewing understandable examples of piping configurations. The cases cited appear to be more applicable to stationary practice, except for the simple loops as shown in Figs. 8 and 9, which are more frequently used in marine piping to provide the required flexibility. It is interesting to note that the simple expansion loop in bending does not create an undesirable amount of localized plastic strains.

In view of the fact that certain pipe-line configurations may increase excessive relaxation stresses which could cause failure as a result of localized concentrations of plastic strains, the maximum permissible limit of the relaxation stress should be indicated in the Code for the guidance of the pipe stress-analysis engineer. Since the Code (B31) has established the stress range as the primary criterion of flexibility of the piping system, in lieu of the individual stresses at any one time during the temperature cycle, the information as to the limiting relaxation stress appears to be lacking. Under the proposed Rules, the relaxation stress (and

¹¹ Head Marine Engineer, Scientific and Performance Section, Bureau of Ships, Navy Department, Washington, D. C. Mem. ASME.

¹² Section Engineer, Mechanical Development Division, The M. W. Kellogg Company, New York, N. Y. Mem. ASME.

¹³ Captain, U. S. Coast Guard, Assistant Chief, Office of Merchant Marine Safety, U. S. Coast Guard Headquarters, Washington, D. C.

the resultant plastic strains) is considered to affect only the strength of the system at the terminal points.

It should be recognized that a companion factor to excessive localized plastic strains which could induce failure is the old culprit—stress concentrations due to poor welding, notches, and sudden changes in metal thickness. As the recent Rankin and Clark paper so clearly pointed out, general yielding of the pipe under temperature cycling should be considered an important factor in preventing fatigue or stress-corrosion cracking; however, points at which yielding cannot occur, such as stress concentrations, should be eliminated from the system, if failure is to be avoided.

S. W. SPIELVOGEL.¹⁴ In this stimulating paper the author presents a novel approach to the evaluation of the residual stresses in pipe loops which on further development should become a useful tool in forecasting the changes in configuration and life expectancy of piping systems.

The Task Force on Flexibility for the Code of Pressure Piping was confronted with this very problem.

In the absence of sufficient analytical research, however, the committee confined its recognition of the phenomenon of self-springing to the establishment of rules for recording the reactions. The retention of partial credit for reactions is due primarily to the uncertainty that the specified cold spring is actually realized because the majority of pipe installations lack the strict supervision necessary to meet the refinements of an exact calculation.

The author in his introduction states that the new issue of the Code implies that there is no particular advantage in cold springing. This interpretation probably is prompted by the statement in the preamble:

"Inasmuch as the life of a system under cyclic condition depends on the stress range rather than the stress level at any one time, no credit for cold spring is warranted with regard to stresses."

In defense of this opinion one must admit that credit allowances for cold spring lose significance unless it is possible to appraise the degree of relaxation at specific temperatures and calculate the resulting residual stress at the sustaining level.

Such information is now being made available by the author's most interesting paper.

The theory which he develops shows that the residual stresses in a pipe loop connected to unyielding anchors can be determined in the same manner as the stresses in a flange bolt in which the initial take-up remained unchanged. Under this condition the sum of the elastic and plastic extensions remains constant. As the former decreases with time the latter increases without change in total deformation. The stress to which the bolt will adjust itself with time can be calculated from Equation [4] of the paper, provided the creep stress for the nominal creep rate and the slope of the log-log plot of tests are known.

A study of the geometry and flexibility of the elements of pipe loops leads to a similar expression for the residual stress during self-springing. The examples of Figs. 8 and 9 give a vivid demonstration of the unsuspected behavior of two loops of identical flexibility. For operation in the elastic range one loop would be as acceptable as the other. In the plastic range, however, the analysis reveals differences in creep and final stress. Although these differences in the particular examples are not strikingly large the theory proves the tendency and the possibility that greater variations may be found on further exploration of more complicated configurations.

The example of the multiple loops of Fig. 10, apparently chosen for its academic interest and with a look into the future when much higher operating temperatures will call for the extreme in

flexibility, shows the surprising tendency for confining plastic strains to a limited region.

In the light of this paper this writer admits that the advantage of 100 per cent cold springing deserves to be emphasized.

The theory developed is indeed a valuable contribution to the art. It stimulates the imagination and points the direction in which to travel but in order to put the result in the hands of the practicing engineer not always familiar with the ramifications of the theoretical background, a condensation to the essentials and a specific working procedure should be the next step, to spread interest and invite verification of the predictable displacements in the field.

J. D. CONRAD.¹⁵ Mr. Robinson capably presents a relatively new factor to be considered by high-temperature piping designers. As stated in the paper, the wording of the revised Pressure Piping Code contains a suitable warning on this matter.

The writer would like to point out, however, that while excessive creep is a factor worthy of consideration, sufficient evidence does not appear in piping experience or design analyses to justify an overnight re-evaluation of all piping installations.

At the time the B-31.1 Task Force on Flexibility was considering this matter a creep analysis of a simple U bend with square corners was made. The results of this analysis are given in Fig. 17 of this discussion.

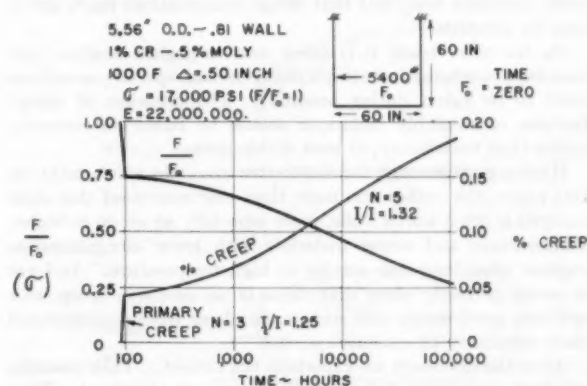


FIG. 17

Note that for this case the predicted creep is 0.18 per cent in 100,000 hr. The most severe case given in the paper appears to be the one shown in Fig. 11 for which the predicted creep in 10,000 hr is 0.26 per cent. An extrapolation of this parallel to Fig. 17 predicts 0.39 per cent in 100,000 hr for Fig. 11. These amounts of creep are worthy of the designer's consideration in some materials of low ductility but surely would not be alarming in most materials used in pipe lines today.

The panacea of cold spring has merit. But again this must be used with judgment and careful supervision by the pipe-line erector. For instance, a U bend can be cut short 100 per cent and erected by the use of a vertical and a horizontal cable to cold-pull an end point, thus applying no corrective moment. For the U bend shown in this discussor's Fig. 17 and using such a cold-spring technique, the weld gap will be decreased less than $1/32$ in. on the inside and increased less than $1/32$ in. on the outside of the pipe bend, a variation too subtle for most erectors with present-day techniques. Yet in this case the hot elastic stresses in the system would not be reduced at all at some locations.

Of course, the closing weld can be made at a point which re-

¹⁴ Consolidated Edison Company of New York, Inc., New York, N. Y.

¹⁵ Manager, Engineering Services Division, Steam Division, Westinghouse Electric Corporation, Philadelphia, Pa. Mem. ASME.

quires no corrective cold moment. In three-plane systems it is not always possible to locate such a point, let alone get at it physically in the field. Or, pipe lines could be erected with strain gages utilized to measure the actual setup stress and thus the effective cold spring in the system. If the proper setup stress were not attained, then the pipe would have to be cut and rewelded till correct. These points are discussed to emphasize the fact that if concentrated creep is a serious problem, the use of cold spring (in lieu of redesign) to counteract the creep demands equally serious considerations. Cold-spring techniques employed in the past have not always been compatible with such careful design techniques.

AUTHOR'S CLOSURE

The author is grateful to all of the discussers for saying the things which needed to be said about this paper and which could be said more effectively by independent experts than by the author himself. In presenting this paper he made the remark that "This analysis does not show that all existing pipe lines are no good. If it did, you would say right off that my analysis is no good."

(For 15 years this problem was always in the back of my mind and it was only the spur of the revision of the Code for Pressure Piping that caused me to write the paper at this time. Throughout this whole period I had felt, like some of the discussers, that most pipe lines work and that design computations must, therefore, be adequate.)

On the other hand, it is never wise to neglect analysis just because of mathematical complexity or because special precautions need to be taken during assembly. The omission of design features or assembly technique should be based on certainty rather than uncertainty, at least within reason.

Having gone through the illustrative examples which make up this paper, the author was more than ever convinced that such analysis is often worth while, more especially as we go to higher temperatures and newer materials with lower elongations-to-rupture after long-time service at high temperature. And yet it seems perfectly plain that there is no excessive creep with ordinary good design and such materials as have demonstrated their suitability by extensive service.

All of the discussers are experts in the subject. Their remarks are both thoughtful and sympathetic as well as critical. Thus their comments should be of great value in enabling the reader to assess the importance of employing some such procedure as here suggested in connection with any particular case of high-temperature piping design and assembly.

Mr. Markl properly notes that the criterion for damage must vary with materials. The author concurs that this is most important, especially when higher-strength materials are accompanied by lower available long-time elongations. In his mathematical analysis Mr. Markl uses the same line of reasoning as used by the author in an earlier paper.¹⁶ Mr. Markl's *damage factor f* is identical with what the author called in his earlier paper *fractional expenditure of life E* . Without going through the arithmetic in detail, the author sees no reason to question the outcome. What he does wonder is this: Are we looking at what is of most importance?

¹⁶ "Effect of Temperature Variation on the Long-Time Rupture Strength of Steels" by Ernest L. Robinson, Trans. ASME, vol. 74, 1952, pp. 777-781.

Long-time rupture testing is mostly constant-stress testing. The author agrees with Mr. Markl that it is hard to see how relaxation with limited total extension can lead to rupture at all. However, when the design is such that widely distributed elastic strain keeps aggravating the plastic extension in some local region, then what is to be expected? A rational determination of the fractional expenditure of life or damage factor is one legitimate way of evaluating a design. In the present paper the author has chosen to evaluate the extent of plastic strain in the worst location as of greatest direct significance when, as Mr. Markl says, "good initial ductility . . . may be greatly reduced under prolonged heating."

Mr. Michel's comments are sympathetic and the author is grateful.

Mr. Murphy restates the objective of this paper very well. He makes out a good case for the harmlessness of self-springing. He likens it to a hot-forming operation. This gets very close to the crux of the matter. Indeed, the author has often wondered if hot-forming operations should not be subject to more critical control. He suggests that many obscure difficulties might be better understood if the strains associated with the relief of hot-forming stresses were better analyzed. For all its complexities, a piping system may be easier to analyze than what goes on inside a chunk of metal during hot-forming.

Captain Smyth points out that effective application of the proposed analysis would require the naming of allowable relaxation stresses by the Code authority. The author is not sure that data are available on all materials to implement such action but, at least, the Piping Code Committee would be no worse off in this respect than the ASME Boiler Code Committee with respect to many of the materials for which it is regularly asked to choose stress allowances.

Mr. Spielvogel explains the implications of the Code wording and welcomes the theory presented but points out the need for a working procedure. With all of this the author agrees. As for implementing the Code, this matter properly belongs in the hands of the Code Committee. The author wonders if it would not be best, for the present at least, to leave the matter of plastic analysis to the discretion of the responsible engineer. However, without suggesting it be mandatory, he would be glad to see any direct or implied penalty on cold-springing or any discounting of its potential advantages eliminated from the wording because this is one of the most effective ways of reducing plastic deformation.

Both Captain Smyth and Mr. Spielvogel have called attention to further work that needs to be done. Maybe the author will get to it some day. In the meantime he would be glad if anybody else would feel like carrying on from here.

Mr. Conrad concludes that creep in a simple loop of ordinary material is not likely to be severe and with this the author agrees. Indeed, it is because of this that he put off writing this paper for so long a time. Mr. Conrad gives a good discussion of the difficulties which sometimes go with effective cold-springing. However, as he said before, the author believes that the only good justification for the neglect of difficulties is the assurance that there will be no seriously adverse consequences.

It seems plain that procedure that is usually satisfactory is not necessarily always so and that the need for good plastic analysis of high-temperature piping is likely to increase in the future rather than disappear.

Stack Heights Required to Minimize Ground Concentrations

By E. W. HEWSON,¹ ANN ARBOR, MICH.

Aerodynamic and meteorological concepts are combined in a procedure for estimating ground concentrations of effluents from stacks with various possible heights and exit gas velocities. The operation of each of the several influences at work is first described. A detailed example is then given of how the most important phases were integrated into a consistent procedure for predicting ground concentrations in answer to a specific design problem. Further improvements and refinements of the method are desirable and are being incorporated in a later study.

INTRODUCTION

THE atmosphere is a vast reservoir which may be used for the disposal of industrial wastes. Used wisely, it permits disposal without damage or nuisance; used without due consideration of its widely varying diffusing capacity, wastes may at times remain at sufficiently high concentrations near the ground to cause annoyance, and even illness and death.

A primary factor in determining whether or not industrial wastes emitted to the atmosphere will be a nuisance is the stack design. Generally speaking, the higher the stack and the greater the exit velocity the better. But improvement at the ground does not vary proportionally with increasing values of these, and economic and other factors limit stack heights and exit velocities. There are thus optimum design values which vary with climatic conditions, topography, plant configuration, composition, temperature, and exit velocity of effluents, and so on. The purpose of the present paper is to suggest to the stack designer a method of approach which will permit the computation of the concentrations to be expected with various stack characteristics over level terrain. In rough country and near shorelines modifications of the method are required.

THE EFFECTIVE STACK HEIGHT

Because of the several influences acting on the effluent gases, the latter often stream downwind with their center lines at a height substantially different from that of the mouth of the stack. The height at which the gases level off has been called the "effective stack height" by N. R. Beers (1).² The first part of the problem requiring consideration is how to determine this effective stack height. Wind-tunnel studies, such as those made by R. H. Sherlock and E. A. Stalker (2, 3), offer an answer. The technique is briefly as follows: A model of the stack and nearby buildings is placed in the tunnel and the smoke emitted from the stack is photographed and its height in the tunnel thereby determined; the influence of varying wind speed is simulated by changing the rate of smoke emission from the stack, tunnel air

speed remaining constant; the model is rotated in the tunnel to simulate various wind directions, and the photography repeated for each wind direction. On the basis of a number of actual model studies, R. H. Sherlock and E. J. Lesher (4, 5) have generalized the method so as to permit, when there are no unusual conditions present, estimates of the behavior of the smoke plume without the necessity of making wind-tunnel studies.

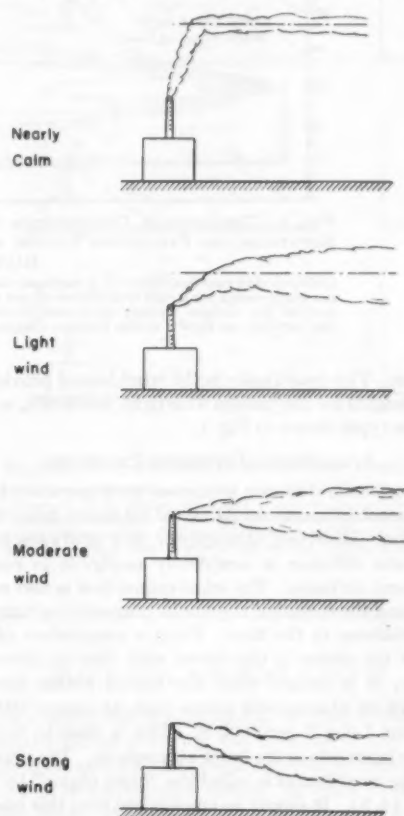


FIG. 1 VARIATION OF EFFECTIVE STACK HEIGHT WITH WIND SPEED

Let us consider the behavior of the stack gases as wind speed increases. With nearly calm conditions the gases rise almost vertically; high stack-gas temperatures and high velocities of emission may cause the gases to rise to considerable heights; ascent to an effective stack height twice the actual stack height is not uncommon. As the wind speed increases the height at which the plume levels off becomes progressively lower. The effect is illustrated in Fig. 1. For a moderate wind the effective stack height is very little greater than the actual stack height. With stronger winds aerodynamic down-wash of the gases may occur. If the stack is on a large building or near a complex of structures the down-wash may be very pronounced. It is difficult to simulate nearly calm conditions in the wind tunnel; for these the effective stack height was obtained by another means, which will be de-

¹ Professor of Meteorology, Department of Civil Engineering, University of Michigan.

² Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Power Division and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 11, 1954. Paper No. 54-A-211.

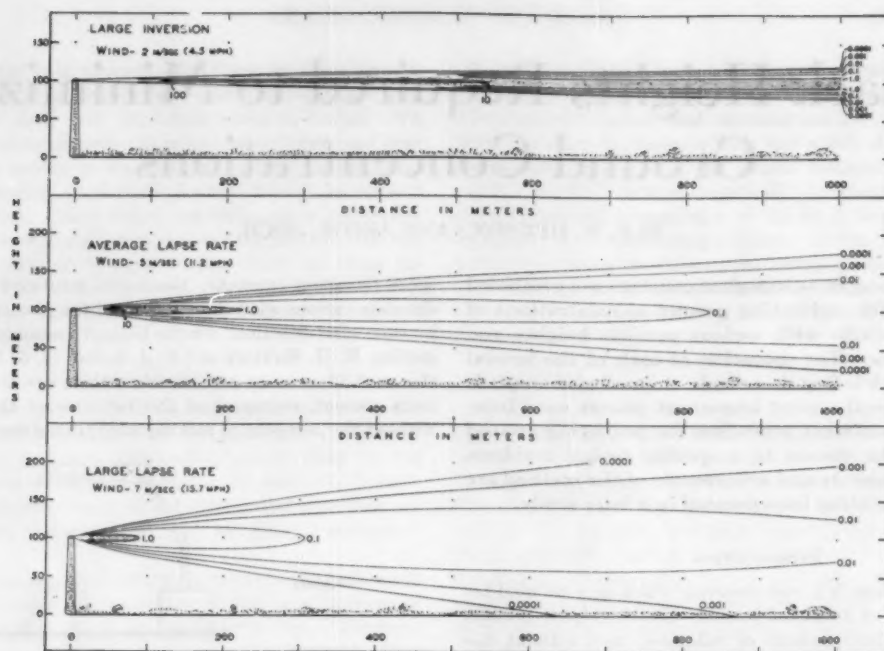


FIG. 2 THEORETICAL DISTRIBUTION OF POLLUTION FROM A 100-METER STACK, FROM SUTTON'S EQUATIONS AND PARAMETER VALUES, FOR THREE CONDITIONS: LARGE INVERSION, AVERAGE LAPSE RATE, AND LARGE LAPSE RATE

(Isolines give concentrations of a contaminant in the vertical plane through the center of the plume in milligrams per cubic meter for a rate of emission of one gram per second. The ground presents a barrier to downward diffusion at the surface leading to increased concentrations there and, hence, to an asymmetrical distribution in the vertical, as shown in the bottom diagram. For simplicity, the effective stack height is assumed equal to the actual stack height.)

scribed later. The model tests in the wind tunnel provided effective stack heights for the plumes with light, moderate, and strong winds of the types shown in Fig. 1.

ATMOSPHERIC DIFFUSION PROCESSES

If only molecular diffusion processes were operative in the atmosphere, stack effluents would travel for many miles with very little dilution. However, atmospheric flow is always turbulent, and molecular diffusion is completely negligible in comparison with turbulent diffusion. The wind-tunnel flow is also somewhat turbulent, and the spread of the smoke plume in the tunnel is due to this turbulence in the flow. From a comparison of the appearance of the plume in the tunnel with that of plumes in the atmosphere, it is judged that the tunnel plume corresponds roughly with an atmospheric plume such as occurs with a lapse rate of about 5 deg F per 1000 ft. This is close to the average value of the lapse rate in the lower atmosphere. The plume which occurs in the wind tunnel is called the "basic plume" by Sherlock and Leshner (4, 5). It should be emphasized that this basic plume will vary slightly in dimensions from tunnel to tunnel, depending on the scale of the turbulence generated in each.

General equations expressing the mean concentration of a contaminant downwind from a continuous point source of emission have been given by C. H. Bosanquet and J. L. Pearson (6) and by O. G. Sutton (7). The latter's expression has been found convenient to use in the present study. In functional form Sutton's equation is

$$\chi = \phi(Q, H_e, v_w, x, y, z, n, C_y, C_z)$$

or in explicit form

$$\chi = \frac{Q \exp \left(-\frac{y^2}{C_y^2 x^{2-n}} \right)}{\pi C_y C_z v_w x^{2-n}} \left[\exp \left(-\frac{(z-H_e)^2}{C_z^2 x^{2-n}} \right) + \exp \left(-\frac{(z+H_e)^2}{C_z^2 x^{2-n}} \right) \right] \quad \dots \dots [1]$$

† $\exp -k$ is equivalent to e^{-k} .

If the contaminant is sulphur dioxide, then in this equation

χ = mean concentration in mass of SO_2 per unit volume of air

Q = rate at which SO_2 is leaving stack expressed as mass per unit time

H_e = height above ground of center of horizontal portion of plume, i.e., effective stack height

v_w = mean wind velocity

x, y, z = co-ordinates along wind, horizontally across wind, and vertically upward, respectively, of point for which concentration is being determined, origin being at ground level at base of stack

n = a diffusion parameter determined from vertical wind profile

C_y, C_z = virtual diffusion coefficients for diffusion in y and z directions, respectively.

Vertical cross sections of the plume, obtained from Equation [1] with $y = 0$, i.e., along the center line of the plume, are shown for three meteorological conditions in Fig. 2. For simplicity the actual and effective stack heights are taken to be equal. The values of the diffusion parameters given by Sutton (7) have been used in the computations. When air temperature increases with height, i.e., in an inversion, an atmospheric condition most prevalent at night, winds tend to be light and turbulent diffusion weak. As the top diagram of Fig. 2 suggests, under these conditions the smoke flows for many miles in a thin but gradually widening band; no appreciable amount of smoke reaches the surface, even at great distances away. The bottom diagram of Fig. 2 gives concentrations when there is a large lapse rate of temperature. The term "lapse rate" signifies the rate at which temperature falls off (lapses) with increasing height in the atmosphere. Large lapse rates occur most frequently on clear afternoons after the ground has been strongly heated by the sun. In such conditions winds are generally moderate and turbulent dif-

fusion is pronounced; concentrations fall off rapidly with distance downwind. The middle diagram represents an intermediate condition.

Maximum ground concentrations occur beneath the center line of the plume where $y = z = 0$, for which Equation [1] simplifies to the form

$$\chi' = \frac{2Q}{\pi C_y C_p x^{3/2}} \exp - \frac{H_s^2}{C_y^2 x^{3/2}} \dots \dots \dots [2]$$

This is a convenient equation for the computation of ground-level concentrations. Detailed techniques of using it in design problems will be described in a later section. A comprehensive discussion of these equations and their significance has been given by Hewson (8).

INVERSION BREAKUP PROCESS

Center-line concentrations at the ground are obtained with satisfactory accuracy from Equation [2] under nearly all meteorological conditions. A notable exception, however, occurring with the inversion breakup, was first described and explained by Hewson (9). During the early morning hours, before sunrise, inversion conditions often prevail, and the plume travels downwind for many miles as a thin but gradually widening ribbon of smoke. After sunrise in summer, solar radiation heats the earth's surface strongly, which in turn heats the air just above, producing a layer with a large lapse rate and pronounced turbulence. As the morning progresses and solar heating of the ground increases, the top of this heated turbulent layer grows steadily upward. When it reaches the ribbon of highly concentrated smoke aloft the latter is rapidly diffused downward to the ground, producing high concentrations there very suddenly. The several stages of the breakup process are illustrated in Fig. 3. As the highly turbulent layer continues to grow, upward diffusion occurs and lateral diffusion becomes increasingly effective. The result is that surface concentrations fall off exponentially with time from their peak values, as shown by the series of graphs appearing at the right of Fig. 3.

The breakup condition is of great importance when high surface concentrations, even if of short duration, must be considered. According to Lowry (10), experiments at Brookhaven National Laboratory show that the average surface concentration during the 15-min period of a breakup with peak values is 20 times the maximum specified by Sutton's equations.

An equation for the average surface concentration occurring during the peak of a breakup is developed in the following. The general outlines of the procedure were described by M. E. Smith (11). During inversion conditions with light winds the plume from a stack flows downwind with sufficient horizontal diffusion to produce a lateral spreading with an angle of approximately 5 deg. The vertical diffusion is extremely small, much less than the horizontal diffusion. Let us consider the mean concentration of sulphur dioxide, for example, at the surface below a plume making a horizontal angle of 5 deg and for the area bounded by arcs of two circles of radii x_2 and x_1 centered at the stack and by the two radii subtending an angle of 5 deg at the stack. The mass of SO_2 vertically above this area is given by Qt , where Q is the stack output in mass of SO_2 per unit time, and t is the time required for the wind of velocity v_w to carry the gas from x_1 to x_2 , or $(x_2 - x_1)/v_w$. As the ribbon of SO_2 is diffused downward, after the upward growing turbulent layer has reached the ribbon, as depicted in Fig. 3, the gas comes to the surface in a concentration which is assumed to be the mean concentration in the volume between the plume and the ground in the distance interval between x_2 and x_1 . The volume below the plume is $\pi H_s(x_2^2 - x_1^2)/72$, where H_s , the effective stack height, is with sufficient accuracy the height of the top of the plume just before its breakup, and the mass of SO_2 in

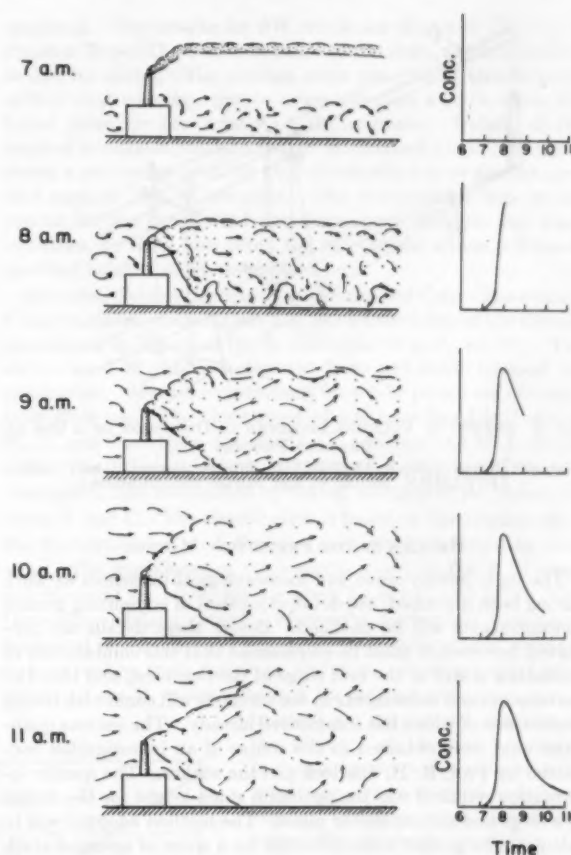


FIG. 3 PROCESS OF DEVELOPMENT OF BREAKUP CONCENTRATIONS (As turbulence grows upward on a clear summer morning, gas is suddenly diffused downward from plume aloft and produces high concentrations at surface. Resulting variation of ground concentrations during morning is sketched on graphs at right. Note sudden increase followed by exponential decrease, which produces very high values during period of approximately $1/2$ hr.)

this volume is $\pi \chi H_s(x_2^2 - x_1^2)/72$. But the two masses are equal, so that

$$Q(x_2 - x_1)/v_w = \pi \chi H_s(x_2^2 - x_1^2)/72$$

However, since there is some lateral diffusion of the SO_2 as well as downward diffusion during the breakup of the plume, approximate allowance for this lateral diffusion is made by doubling the volume in which diffusion occurs. In order to conform with customary usage, the concentration is expressed by parts per million (ppm) by volume and denoted by X . With these provisions

$$X = 36Q/a\pi v_w H_s(x_2 + x_1) \dots \dots \dots [3]$$

where a is a conversion factor used in changing to ppm.

The model assumed is shown in Fig. 4. The distance from the center line of the stack to the surface $ABCD$ is x_1 ; that to the surface $EFGH$ is x_2 . Height H_s is AJ or EP . During the breakup of the inversion the SO_2 initially in volume $ABFEDCGH$ diffuses vertically downward into volume $ABFEJKQP$, or if horizontal diffusion is considered as well, into volume $ABFELMRN$, which is twice as large, leading to the average concentration X as specified by Equation [3]. According to M. E. Smith (11), a formula of this type has been applied to a few oil-fog tests at Brookhaven National Laboratory with fair success.

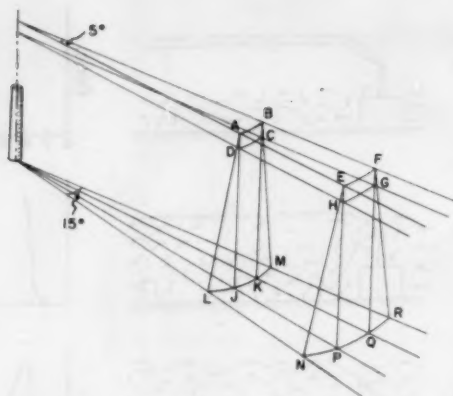


FIG. 4 SKETCH OF VOLUMES INVOLVED IN DIFFUSION OF A GAS BY BREAKUP PROCESS
(Gas initially in volume ABFEDCGH diffuses downward into volume ABFELMRN, leading to high surface concentrations.)

DETAILS OF THE PREDICTION METHOD

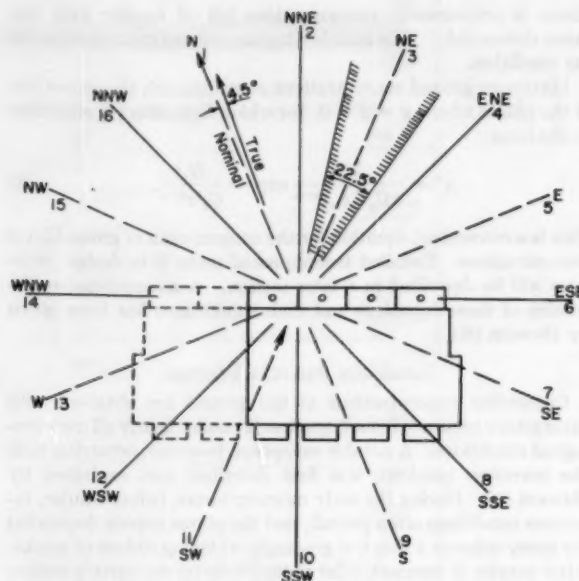
The basic aerodynamic and meteorological influences at work having been described, the detailed method of computing ground concentrations will be outlined. Before these details are presented, however, it must be emphasized that this combination of specialties is still in the first stage of development, and that improvements and refinements in the methods will come with future studies, one of which has commenced already. The various quantities used were obtained in the course of an investigation conducted by Prof. R. H. Sherlock and the author. The specific information required was the optimum stack height for the design of a large midwestern power plant. The method adopted was to calculate the ground concentrations for a series of assumed stack heights. On the basis of the concentration values supplied, a stack height was selected which was considered to be the optimum in the light of all the various factors involved.

Obtaining Annual Duration of Aerodynamic Down-Wash. The orientation, size, and proportions of the projected plant were obtained from the preliminary drawings furnished. The axes of the plant do not correspond exactly to any of the cardinal directions of the compass or the subdivisions of them, as used by the U. S. Weather Bureau. In order to simplify the computations it was assumed that the axis of the line of stacks lay on a line which runs from ESE to WNW. This is in error by about 4.5 deg. Fig. 5 shows the adopted or nominal directions as compared to the true directions. It will be noticed that the area around the plant is divided into 16 segmental parts each having an included angle of 22.5 deg.

It is first necessary to compare the size and proportions of the projected plant with those of the hypothetical plant described by Sherlock and Lesher (4, 5). The critical conditions for aerodynamic down-wash around the latter plant had been determined from wind-tunnel studies. In all the principal dimensions the new plant is larger than the hypothetical plant, the greatest difference being in height, 180 ft compared to 125 ft, giving a ratio of 1.44. Corresponding ratios are 1.21 for width and 1.09 for length.

The heights H of stacks for the hypothetical model were 250, 300, 350, or 400 ft and the height h of the building was 125 ft, giving values for the ratio H/h of 2.0, 2.4, 2.8, and 3.2. Multiplying these model stack heights by 1.44 gives corresponding stack heights of the new plant of 360, 432, 504, and 576 ft.

Wind-tunnel studies on the hypothetical model established the air speeds which are necessary to maintain the bottom of the plume at various heights P above the ground, namely 100, 150,



NOMINAL ORIENTATION FOR THIS PROJECT

Example: SW winds blow from the SW and affect the 22.5° segmental area as shown cross-hatched.

FIG. 5 DETAILS OF PROPORTIONS AND ORIENTATION OF NEW PLANT AND POSSIBLE LATER ADDITIONS, LATTER SHOWN BY BROKEN LINES
(Winds classed as SW will carry smoke over segmental area of 22.5 deg to NE of plant outlined by hatching and similarly for other directions.)

200, and 250 ft. Multiplying by 1.44 to correct for the greater height of the new plant gives values of P of 144, 216, 288, and 360 ft. For convenience these were rounded off to 140, 215, 285, and 360 ft. The ratio of stack-gas velocity to wind velocity which brings the base of the plume to any specified height P is referred to as the "critical velocity ratio," v_s/v_w .

As an example of the method, the number of hours per year of aerodynamic down-wash from a stack 360 ft high which brings the base of the plume to 140 ft above the ground with southwest winds will be calculated. The height of the new plant is 180 ft, that of stack is 360 ft, so that ratio $H/h = 2.0$. The critical velocity ratio for $P = 140$ ft is obtained from Fig. 6 as given by Sherlock and Lesher (4, 5). For a SW wind (direction 11) and a value of P of 140 ft (corresponding to a value of P of 100 ft with the hypothetical plant), v_s/v_w is read off to be 2.3. The stack velocity is 120 fps at a temperature of 250 F. In comparisons involving the momentum of stack and model effluents the differences in density should be allowed for: The density of the stack effluent is smaller than that of the model effluent since the temperature of the former is greater—250 F in comparison with 70 F. The equivalent stack velocity at 70 F is given by

$$v_s = 120 \frac{70 + 460}{250 + 460} = 90 \text{ fps} = 61 \text{ mph}$$

Since $v_s/v_w = 2.3$, $v_w = 26.6$ mph. Winds from the southwest with speeds greater than this critical value will, by aerodynamic down-wash, bring the base of the plume down to 140 ft above ground or lower. The next step is to determine how often such winds occur at or near the site of the new plant.

Wind records taken in that area by the U. S. Weather Bureau during the period August, 1949, to July, 1953, inclusive were

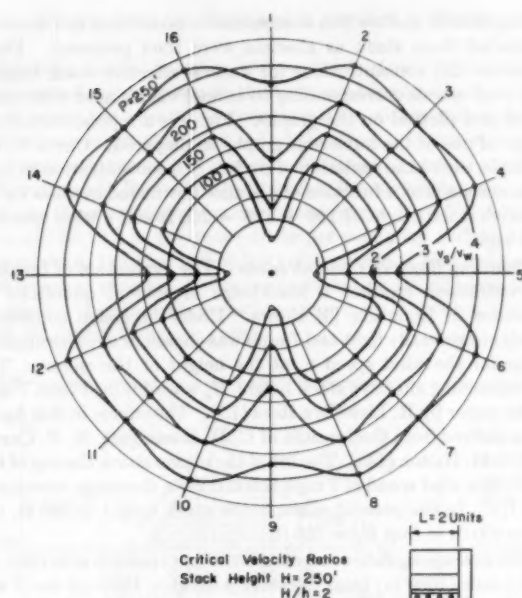


FIG. 6 POLAR DIAGRAM GIVING CRITICAL VELOCITY RATIOS v_c/v_w FOR THE HYPOTHETICAL MODEL PLANT AS A FUNCTION OF WIND DIRECTION

(After R. H. Sherlock and E. J. Leshner.)

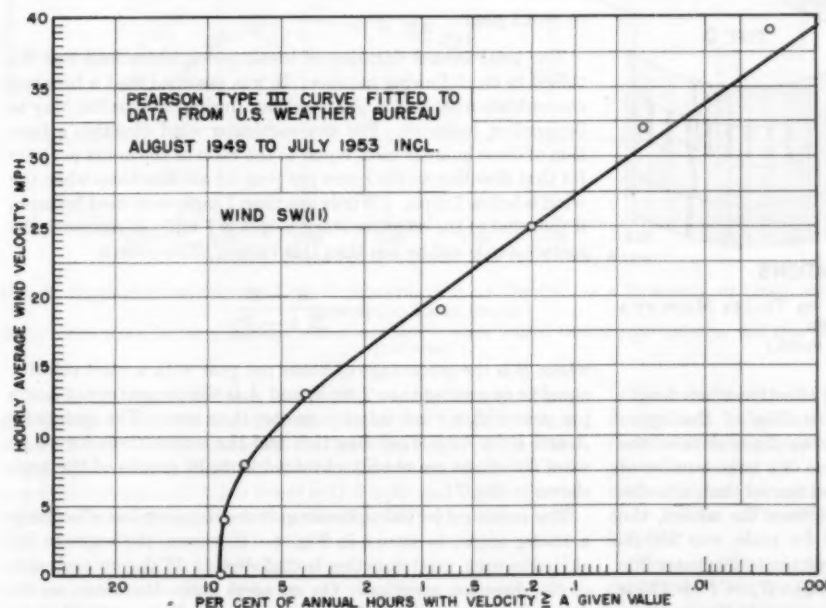


FIG. 7 PERCENTAGE FREQUENCY OF OCCURRENCE OF SW WINDS WITH VELOCITIES EQUAL TO OR GREATER THAN SPECIFIED VALUES

(U. S. Weather Bureau data, August, 1949, to July, 1953, inclusive.)

analyzed. The results for SW winds are shown in Fig. 7. A Pearson Type III curve was fitted to the data, which are represented by circles. The abscissa gives per cent of annual hours with a wind velocity equal to or greater than a given value, the latter given by the figures on the ordinate. Details of the method have been published by R. H. Sherlock (12). The figure shows a percentage of 0.118 with speeds equal to or greater than 26.6 mph or 10.3 hr per year. The computation may be repeated for the other wind directions, stack heights, and stack velocities, to show how often the base of the plume will reach specified heights above the ground.

Obtaining Surface Concentrations for General Case. The evaluation of surface concentrations requires a knowledge of the various parameters in Equation [2], in particular of n , C_p , and C_s . The values used in this investigation have not been released for publication. Mean concentrations for a 1-hr period are obtained from their use. The three types of gustiness listed in Table 1, B_2 , B_1 , and C are those described by I. A. Singer and M. E. Smith (13). Type E was devised for the special purposes of this investigation, and is obtained by taking averages of the values for types B_1 and C. The classification is based on the amplitudes of the fluctuations of wind direction as shown by a recording wind vane. The appearance of the traces is shown in Fig. 8. Precise definitions of the categories are given by Singer and Smith (13). Generally speaking, type C represents mechanical turbulence generated in the air as it flows over surface-roughness elements such as bushes, trees, buildings, and so on. It occurs with higher winds. Type B_1 is characterized by thermal turbulence super-

imposed on mechanical turbulence; the presence of thermal turbulence is shown by the wider swings of the wind vane. Such thermal turbulence begins to develop when the air becomes slightly stable or slightly unstable. Type B_2 represents turbulent flow with a considerably greater thermal component which develops as instability increases. Type D occurs when the air is very stable, as in early morning inversions with light winds.

The problem is simplified by expressing the types of gustiness as functions of wind speed and time of day, as shown in Table 1. For winds 0 to 7 mph, all turbulence is assumed to be type B_2 or type D except for $1/2$ hr each day for the 5-month period from May to September, inclusive, when breakup conditions are assumed to occur. Parameter values for type D have not been obtained, since measurable concentration values below the plume could not be found.

A further simplification may be made by expressing the gustiness types as functions of wind speed

TABLE 1 TURBULENCE TYPE AS A FUNCTION OF WIND SPEED AND TIME OF DAY

Range of wind speed, mph	Type of turbulence Day	Type of turbulence Night
0-7	B_2	D
8-18	B_1	C
>18	C	C

TABLE 2 TURBULENCE TYPE AS A FUNCTION OF WIND SPEED

Range of wind speed, mph	Type of turbulence
0-7	B_2
8-18	E
>18	C

alone, as is done in Table 2. Such a simplification is not recommended, however. The analysis is facilitated by assuming a type of gustiness, E , with parameters having values which are the mean of those for types B_1 and C . For winds 8 to 18 mph, as Table 1 shows, turbulence is taken to be type B_1 during the day and type C at night. If no distinction between day and night is to be made, an approximation to the values of the parameters appropriate for 8 to 18 mph is obtained by taking the mean of those for day and night. This device may lead to unwarranted oversimplification of the problem, and should not be used except in an emergency. Finally, it must be emphasized that type E has no physical reality in the sense that the other types have, but is merely a quantity introduced for convenience.

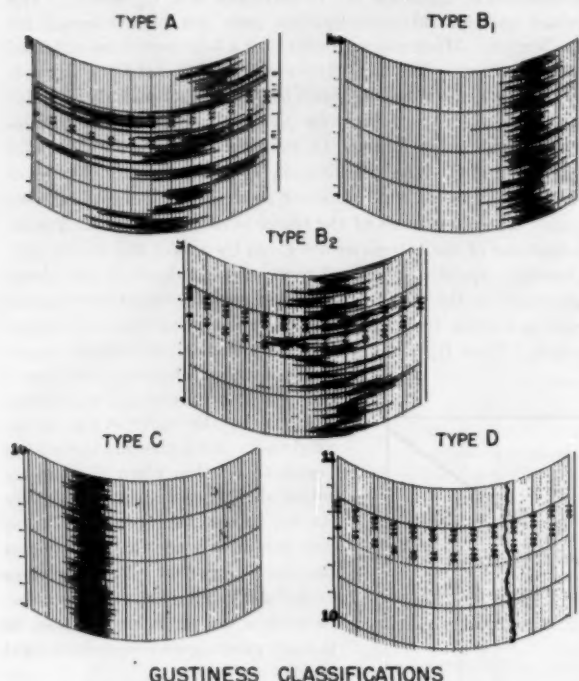


FIG. 8 GUSTINESS CLASSIFICATIONS BASED ON TRACES MADE BY A RECORDING WIND VANE
(After I. A. Singer and M. E. Smith.)

Except during breakup conditions, the effective stack height H_e was obtained from the wind-tunnel studies of the hypothetical plant, but adjusted according to the dimensions of the new plant. It was found from an inspection of a large number of photographs of the basic plume in the wind tunnel that, at a distance of several stack heights downwind from the source, the vertical thickness of the plume, adjusted for scale, was 250 ft. The vertical distance from the base of the plume to its center line is therefore 125 ft. The effective stack height $H_e = P + 125$ is expressed in feet. Where P is 140, as in the example given earlier, $H_e = 140 + 125$ ft.

A rather elaborate procedure has been developed by C. H. Bosanquet, W. F. Carey, and E. M. Halton (14) for the determination of effective stack heights. However, since the method makes no allowance for aerodynamic down-wash, the results are not generally applicable to the new plant. The authors of that paper recommend wind-tunnel studies when large buildings are near the stack. Their results were used, however, in evaluating breakup concentrations, when aerodynamic influences are negligible.

Graphs with surface SO_2 concentration as ordinate and distance downwind from stack as abscissa were then prepared. From Equation [2], concentrations for various effective stack heights and wind speeds corresponding to tunnel values used were computed and plotted on the graphs. Fig. 9 is a representative example of one of the numerous graphs so prepared. Curve R, for example, shows the variation of surface SO_2 concentration in ppm from one stack as a function of distance downwind in miles for an effective stack height of $140 + 125 = 265$ ft and a wind speed of 26.6 mph.

Obtaining Breakup Concentrations. The evaluation of breakup concentrations requires a knowledge of several quantities as reference to Equation [3] shows. Under inversion conditions winds are generally light and 7 mph was chosen as a representative value for the wind speed v_w at the height of the plume. The corresponding effective stack height H_e was obtained from Fig. 6 of the paper by R. Llewellyn Rees (15). The curves in this figure were derived from the formula of C. H. Bosanquet, W. F. Carey, and E. M. Halton (14). The lift of the plume above the top of the stack at a wind speed of 7 mph is 400 ft for a discharge velocity of 120 fps. In the present example the stack height is 360 ft, the lift is 400 ft, so that $H_e = 760$ ft.

The average surface concentration X between 0.5 mile (805 m) and 1 mile (1609 m) from the stack with $Q = 1650$ gm sec⁻¹ will now be computed. The conversion factor a is 2.65×10^{-3} , $v_w = 7$ mph or 3.13 m sec⁻¹, and $H_e = 760$ ft or 231.6 m. From Equation [3]

$$X = 36 \times 1650 / 2.65 \times 10^{-3} \times 3.14 \times 3.13 \times 231.6 (1609 + 805) = 4.1 \text{ ppm}$$

The total annual duration of breakup concentrations was obtained in the following manner: It was assumed that a breakup concentration occurs for $1/3$ hr each day during the months, May to September, inclusive. For any particular wind direction a fraction of these hours is used, equal to the ratio of the hours per year for that direction to the hours per year for all directions when the wind is below 7 mph. (Winds less than 7 mph were used because, if the wind at the effective stack height is 7 mph, as assumed, the surface winds will be less than this value.) This ratio is

$$k = \frac{A - B}{\Sigma(A - B)}$$

where B is the percentage of hours per year with a wind velocity equal to or greater than 7 mph, and A is the percentage of hours per year with a wind velocity greater than zero. The quantities A and B for each wind direction and the summations for all 16 wind directions are readily obtained from 16 graphs of the type shown in Fig. 7.

The sector subjected to breakup concentrations has effectively a 15-deg angle, as shown in Fig. 4. However, the segment included in each wind direction is 22.5 deg, 11.25 deg on each side of the direction specified. On an areal basis, therefore, an individual point in this 22.5-deg segment will be subjected to breakup concentrations for a fraction 15/22.5 of the time.

The number of hours of such concentrations for each wind direction is therefore given by the following expression

$$365 \times \frac{1}{2} \times \frac{5}{12} \times \frac{15}{22.5} k = 51 k$$

Thus, with SW winds, k has the value 0.05, so that there are 3 hr of breakup concentrations per year with SW winds, i.e., to the NE of the plant. The magnitude of the concentrations depends

on the distance from the stack and is specified by Equation [3] for X .

Obtaining Annual Duration of Specified Surface Concentrations. A convenient approximate method was used for finding the number of hours per year when concentrations within stated limits would occur in the neighborhood of the stack. It was shown in the example given in the section, Obtaining Annual Duration of Aerodynamic Down-Wash, in which the quantities used are precisely those employed for computing curve R of Fig. 9, that there will be 10.3 hr per year—10 hr per year with sufficient accuracy—when the wind will be from the southwest with a velocity of 26.6 mph or greater. Reference to Equation [2] will show that this implies that the concentrations shown by this curve will be exceeded during 10 hr per year.

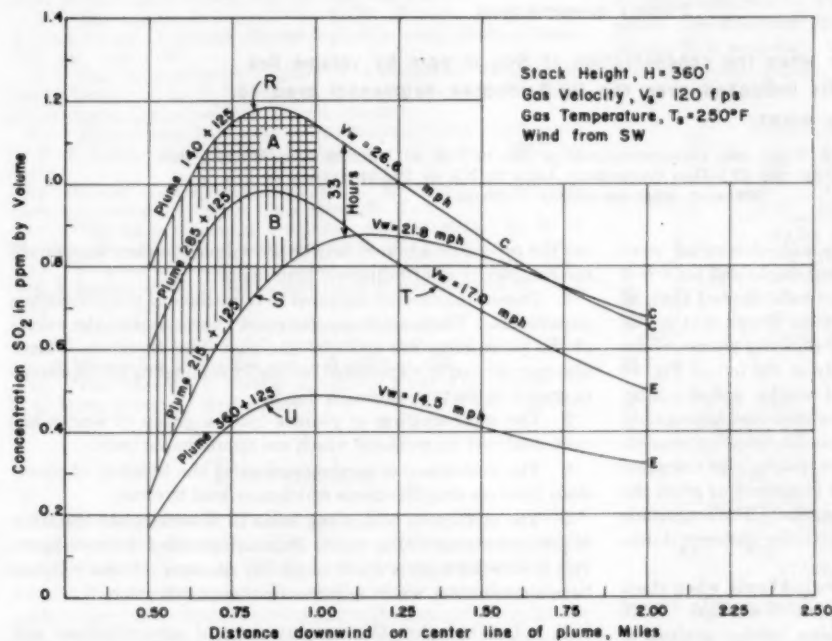


FIG. 9 VARIATION OF CENTER-LINE CONCENTRATION AT GROUND AS A FUNCTION OF DISTANCE DOWNWIND FROM STACK
(Each curve is drawn for a specified plume or effective stack height, critical wind velocity, and type of gustiness.)

Curve S in Fig. 9 is drawn for an effective stack height H_e of $215 + 125 = 340$ ft. A similar analysis of the data shows that when the wind is from the southwest and the height of the stack is 360 ft, the critical wind velocity required to maintain the bottom of the plume at 215 ft or lower is 21.8 mph, and the number of hours per year during which this velocity will be exceeded is 43. This means that there will be 43 hr per year when the surface concentrations will exceed the values shown by curve S. It follows that the number of hours per year when the concentrations will lie between the values given by curves S and R is $43 - 10 = 33$. Similar analyses may be made using curves T and U as well.

The results of the calculations may be conveniently displayed in a diagram such as Fig. 10. In each of the four units of the figure distance downwind from the stack is indicated horizontally and concentration of SO_2 is shown vertically. The stack-gas velocity is 120 fps, the wind is SW, the stack heights are 360, 432, 504, and 576 ft. The values in this array were obtained in the following manner: Consider the left-hand unit of Fig. 10, for a stack height of 360 ft, and take the distance downwind be-

tween 0.5 and 1.0 mile. It will be seen from Fig. 9 that the center-line concentrations during 33 hr of the year, which comprise all concentrations lying between curves R and S, include some concentrations above 1.0 ppm, as shown by area A (cross hatching), and some concentrations below 1.0 ppm, as shown by area B (vertical hatching). The number of hours above 1 ppm is given by $33 [\text{area A}/(\text{area A} + \text{area B})]$. Also, for 10 hr per year, the concentrations will be above curve R, more than 1.0 ppm for approximately 9 hr and less than 1.0 ppm for approximately 1 hr. This division in the ratio of 9 to 1 hr is based on an estimate of the probable areas above curve R which are involved; a more precise evaluation would require a much more refined analysis. The number of hours is thus $33 [\text{area A}/(\text{area A} + \text{area B})] + 9 = 35$ hr. There are thus 35 hr of center-line concentrations between 1 and 2 ppm.

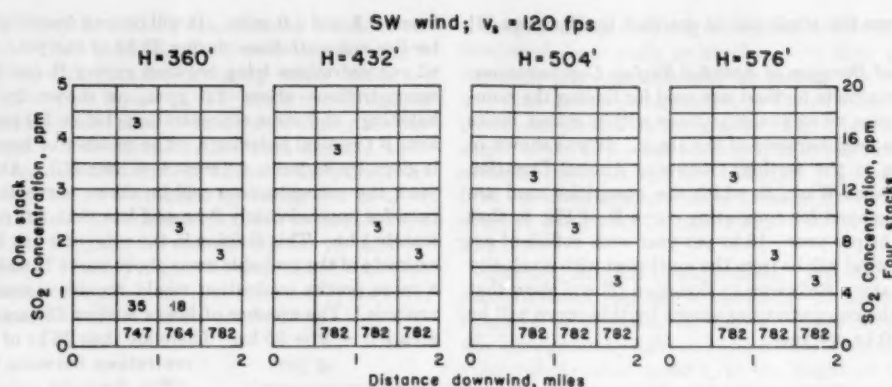
The foregoing values are, however, those on the ground under the center line of the plume. The across-wind values of surface concentration follow an error-function curve, as inspection of Equation [1] shows, with the maximum value vertically beneath the center line of the plume. Large-scale wind-direction fluctuations cause the plume to wave back and forth horizontally, with the result that the concentration at a point varies widely with time. It may be the peak or center-line value at one time and very nearly zero a little later, and so on. For this reason the mean concentration over a segmental area for a single stack is assumed to be one half of the peak concentration. Thus there are 35 hr of concentrations lying, not between 1.0 and 2.0 ppm, but between 0.5 and 1.0 ppm. This value is then entered in Fig. 10 in the appropriate box for concentrations between 0.5 and 1.0 ppm which occur in the distance interval from 0.5 to 1.0 mile.

It was shown in the preceding section that there will be 3 hr of breakup concentrations with SW winds. This value is also entered in each of the appropriate boxes in Fig. 10.

The total number of hours per year with SW winds is readily determined. Fig. 7 shows that the wind will blow from the southwest during approximately 9 per cent of all the hours of the year, i.e., during 785 hr. The number of hours per year with concentrations between 0 and 0.5 ppm is, therefore, $785 - (35 + 3) = 747$ hr. The other values in the array are obtained similarly.

Obtaining Annual Duration of Concentrations With More Than One Stack. The SO_2 concentrations discussed thus far are those on the ground under the plume from a single stack. The method may be extended, however, to cover the case where there is more than one stack. Since the first two installations of the new plant will have a total of four stacks, as shown in Fig. 5, it was necessary to establish a relation between the concentrations produced by a single stack and those caused by four stacks. Such a relationship is readily obtained when, as with the new plant, the distance between the two end stacks is small in comparison with the distances downwind which are of primary interest.

With stacks close together, the concentration under the plume from one stack will be augmented by gases from neighboring



Hours per year when the concentration of SO_2 in ppm by volume lies within the limits indicated over the 22.5-degree segmental area to the NE of the plant.

FIG. 10 HOURS PER YEAR WHEN THE CONCENTRATION OF SO_2 IN PPM BY VOLUME LIES WITHIN THE LIMITS INDICATED OVER THE 22.5-DEG SEGMENTAL AREA TO NE OF PLANT SHOWN IN FIG. 5 (SW wind; stack-gas velocity = 120 fps.)

stacks. Calculations of concentrations a mile downwind were made with Equation [1] for each of the four stacks and for $z = 0$ and $y \neq 0$. Summing the contribution of each showed that, at that distance downwind, an error of less than 10 per cent in the concentrations near the center line of the composite plume results if the concentration values given vertically at the left of Fig. 10 are simply multiplied by the number of stacks, namely, four. The vertical scale to the right in Fig. 10 is therefore appropriate for four stacks. A corresponding analysis for breakup concentrations using Equation [3] was also undertaken, and the same result found to hold, but with a somewhat greater error when the wind blows perpendicularly to the line of stacks. The magnitude of this error decreases, however, with increasing distance downwind from the stacks.

It must be emphasized that these results hold only when stack separation is small compared with distance downwind. With widely spaced stacks or a long line of stacks, further analysis of the problem will be required.

ERRORS DUE TO ASSUMPTIONS

In assessing the reliability of the method described it is desirable to keep in mind the assumptions involved. The obvious assumptions are as follows:

1 The buoyancy effects of hot stack gases are neglected. The model stack effluent is not hot but at the temperature of the tunnel air. The lifting of the plume by jet action is, therefore, simulated but not that arising from the buoyancy of hot stack gases.

2 The effects of mechanical turbulence produced by large stacks and buildings on diffusion downwind from them are neglected. The experimental diffusion parameters used were obtained downwind from a tall slender tower and stack, not downwind from bulky stacks and buildings, which cause extensive and pronounced turbulence in their wake.

3 The influence of atmospheric stability on diffusion is introduced only indirectly. The diffusion parameters were evaluated in terms of types of turbulence, which were in turn related to wind-velocity categories. In particular, the influence of large-scale convective cellular motion on plume behavior and concentrations in unstable air may be appreciable and may require special analysis. Stability and wind velocity are interrelated,

but the procedure adopted may be too simple to reflect adequately the complexity of atmospheric turbulence.

4 Discontinuities are assumed in the values of the turbulence parameters. These arbitrary "stepped" variations in the values of the parameters are artificial and may lead to errors. They also may be partly responsible for the intermingling of the curves in graphs of the type shown in Fig. 9.

5 The intermingling of plumes from a group of stacks has been analyzed by methods which are approximate only.

6 The evaluation of concentrations at the breakup of inversions involves simplifications which may lead to error.

7 The method of comparing areas to determine the duration of concentrations within stated limits at specified distance intervals downwind from a stack implicitly assumes a linear relation between variables which requires further justification.

It is believed that the computations of concentrations and their frequencies are not so greatly in error as to invalidate seriously the predictions of comparative benefits to be derived from different stack heights and different stack-gas velocities. In view of the uncertainties involved, however, the present analysis must be considered as no more than an approach to a satisfactory solution of a very complex problem. Much refinement of the procedures described is desirable and the author, working in association with Prof. R. H. Sherlock, is engaged in a further study in which, it is hoped, some of these uncertainties will be eliminated.

ACKNOWLEDGMENTS

The methods described in this paper have been worked out in close association with Professor Sherlock, who has contributed greatly to the analysis, much more than it has been possible to mention in the main body of the text. This assistance is gratefully acknowledged.

Very substantial contributions to the study also have been made by Messrs. Maynard E. Smith and Irving A. Singer of Brookhaven National Laboratory, whose counsel has been most helpful and is deeply appreciated.

The mass of detail work required in the analysis has been capably handled by Mr. N. M. Isada assisted by Prof. A. L. Coreuera and Messrs. J. A. Arn and J. F. Asuncion.

BIBLIOGRAPHY

- 1 "Stack Meteorology and Atmospheric Disposal of Radioactive Waste," by N. R. Beers, *Nucleonics*, vol. 4, 1949, pp. 28-38.
- 2 "The Control of Gases in the Wake of Smokestacks," by R. H. Sherlock and E. A. Stalker, *Mechanical Engineering*, vol. 62, 1940, pp. 455-458.
- 3 "A Study of Flow Phenomena in the Wake of Smokestacks," by R. H. Sherlock and E. A. Stalker, University of Michigan Engineering Research Bulletin, No. 29, 1941, 49 pages.
- 4 "Design of Chimneys to Control Down-Wash of Gases," by R. H. Sherlock and E. J. Lesher, *Trans. ASME*, vol. 77, 1955, pp. 1-9.
- 5 "Role of Chimney Design in Dispersion of Waste Gases," by R. H. Sherlock and E. J. Lesher, *Air Repair*, vol. 4, 1954, pp. 65-74.
- 6 "The Spread of Smoke and Gases from Chimneys," by C. H. Bosanquet and J. L. Pearson, *Trans. of the Faraday Society*, vol. 32, 1936, pp. 1249-1264.
- 7 "The Theoretical Distribution of Airborne Pollution From Factory Chimneys," by O. G. Sutton, *Quarterly Journal of the Royal Meteorological Society*, vol. 73, 1947, pp. 426-436.
- 8 "Atmospheric Pollution," by E. W. Hewson, *Compendium of Meteorology*, American Meteorological Society, Boston, Mass., 1951, pp. 1139-1157.
- 9 "The Meteorological Control of Atmospheric Pollution by Heavy Industry," by E. W. Hewson, *Quarterly Journal of the Royal Meteorological Society*, vol. 71, 1945, pp. 266-282.
- 10 "Microclimate Factors in Smoke Pollution From Tall Stacks," by P. H. Lowry, *Meteorological Monographs*, vol. 1, no. 4, 1951, pp. 24-29.
- 11 Private communication from M. E. Smith, Brookhaven National Laboratory, Upton, Long Island, N. Y.
- 12 "Analyzing Winds for Frequency and Duration," by R. H. Sherlock, *Meteorological Monographs*, vol. 1, no. 4, 1951, pp. 42-49.
- 13 "Relation of Gustiness to Other Meteorological Parameters," by I. A. Singer and M. E. Smith, *Journal of Meteorology*, vol. 10, no. 2, 1953, pp. 121-126.
- 14 "Dust Deposition From Chimney Stacks," by C. H. Bosanquet, W. F. Carey, and E. M. Halton, *Proceedings of The Institution of Mechanical Engineers*, vol. 162, 1950, p. 355.
- 15 "The Removal of Oxides of Sulphur From Fuel Gases," by R. L. Rees, *Journal of the Institute of Fuel*, vol. 25, no. 148, 1953, pp. 350-357.

Discussion

M. S. OZKER³ AND P. M. GIEVER.⁴ This paper is an excellent complement to a previous paper, Design of Chimneys to Control Down-Wash of Gases, by Sherlock and Lesher, presented at the Annual Meeting last year. The author of the present paper has made a very worth-while contribution to the study of air-pollution control by developing methods for the evaluation of meteorological factors as they affect the dispersion of effluents from stacks by combining these factors with the results of wind-tunnel studies made under controlled conditions. A procedure is developed for specifying stacks, the effluent of which should contribute little or nothing to the air-pollution problem.

In referring to the determination of the "effective stack height" through wind-tunnel experiments, it is mentioned that the influence of varying wind velocity was simulated by changing the rate of smoke emission from the stack of the plant model while the air velocity in the tunnel was held constant. We believe that this point should be elaborated upon as to its justifiability. Modern power-plant buildings now attain considerable height and the influence of such high-building configurations is known to be the cause of aerodynamic down-wash. Since constant air velocity in the wind tunnel would produce a nonvarying turbulence pattern, it seems unlikely that the effect of varying wind velocity can be attained by merely changing exit velocities.

The computations of ground-level concentrations were made by the use of Sutton's equation. It is worth while to note that, in several localities which we have studied, the measured concen-

trations have been less than the predicted concentrations, often by as much as a factor of 6. It thus can be concluded that under usual atmospheric conditions, which do not include the temperature-inversion conditions, the atmosphere does a better job of dispersion and diffusion than predicted by Sutton's equation. Such a situation is very encouraging.

The "breakup" phenomenon resulting from temperature-inversion conditions and its consequences cannot be overemphasized. The author should be commended for the clear presentation of a complex subject and the concise, useful relationships of various factors involved in predicting ground-level concentrations. It should be pointed out, however, that the results appear to be quite discouraging if one considers the means for circumventing such adverse conditions of probable air pollution. It appears that in order to avert the effects of the temperature inversion phenomenon, stacks would have to be designed to discharge above the inversion level. Since this is beyond all practical limits, it is necessary to accept these short periods of higher ground concentrations.

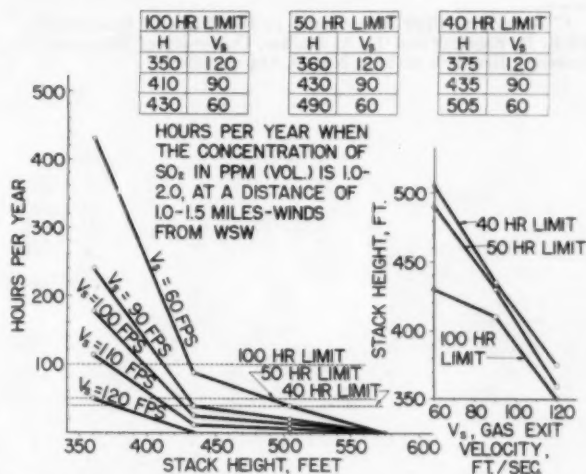


FIG. 11 PREDICTED PERFORMANCE OF VARIOUS STACKS BASED ON AUTHOR'S PROCEDURE

The procedures outlined in the author's paper already have been used in designing new coal-burning power plants. Predicted information can be compiled in the manner shown in Fig. 11 of this discussion, where an attempt is made to indicate equivalent stacks having various stack-height and exit-velocity combinations for various standards of hours per year for a given range of ground concentrations. On the basis of this type of data presentation, economic studies can be made.

It is of interest to note that from the graph on the right of the chart it can be seen that 2 ft in stack height is equivalent to 1 ft per sec in stack-gas exit velocity.

The relationship of climatology to the design and location of stacks has been shown to be of great importance, but another item of equal importance that should not be overlooked is the existing or background level of air pollution in the locality.

Continued observations and sampling of the plume of existing stacks through the use of transits, mobile-testing units, and helicopters as conducted by Thomas,⁵ may produce valuable data for further verifying parameters and formulas now in use. It is believed that more field investigations would be helpful in corre-

³ Assistant Division Supervisor, Mechanical Division, Engineering Laboratory and Research Department, The Detroit Edison Company, Detroit, Mich. Mem. ASME.

⁴ Research Engineer, The Detroit Edison Company, Detroit, Mich.

⁵ "TVA Air Pollution Studies Program," by Fred W. Thomas, Tennessee Valley Authority, Wilson Dam, Ala., *Air Repair*, vol. 4, August, 1954, pp. 7-12.

lating the results of wind-tunnel and mathematical conclusions as used in stack design.

It is concluded that adequate stack design must be accompanied by other necessary air-pollution measures to procure the benefits of design and efforts in order to achieve that elusive objective of adequately clean air.

AUTHOR'S CLOSURE

Messrs. Ozker and Giever ask about the justification of the procedure for obtaining different velocity ratios by varying the stack-gas velocity instead of the wind velocity. Generally, the stack-gas velocity is constant during any particular conditions of plant operation and if there is to be a change in the velocity ratio it will be as a result of a change in the wind velocity. The best justification for the experimental procedures used in the wind tunnel is that predictions based on such wind-tunnel studies have been confirmed by direct observations of the behavior of plumes in the field. The aerodynamic considerations in model testing in the wind tunnel have been discussed earlier in considerable detail.⁶

⁶ "A study of Flow Phenomena in the Wake of Smokestacks," by R. H. Sherlock and E. A. Stalker, University of Michigan, Engineering Research Bulletin No. 29, March, 1941.

It is mentioned that measured concentrations considerably less than values predicted by Sutton's equations have been found. Sutton's published values of the diffusion coefficients were developed on the basis of 3-min samples. If coefficients appropriate for 1-hr samples are used, as in the present paper, then Sutton's equations give concentrations smaller than those presented in his original article.

Regarding breakup concentrations, it is true that higher stacks offer only limited benefits. There are other possible remedies than higher stacks, however, which engineering skills may well develop in the future in, and when, the problem of breakup concentrations becomes acute.

Messrs. Ozker and Giever also present Fig. 11 and point out an equivalence between an increase in stack height and an increase in stack-gas exit velocity. It should be emphasized, however, that such an equivalence varies from one situation to another and that it is invalid to conclude that there is an invariable relation between these two factors.

The author heartily supports the suggestion that more field investigations should be undertaken in order to provide data for improved design practices. Substantial benefits may be expected from relatively modest field programs.



